



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *01/12/2014* par :

LAURE SOULIER

**Définition et évaluation de modèles de recherche d'information
collaborative basés sur les compétences de domaine et les rôles des
utilisateurs**

JURY

FLORENCE SEDES	Professeur, Université de Toulouse 3	Présidente du jury
ERIC GAUSSIER	Professeur, Université de Grenoble 1	Rapporteur
JIAN-YUN NIE	Professeur, Université de Montréal	Rapporteur
MAARTEN DE RIJKE	Professeur, University of Amsterdam	Examineur
PATRICK GALLINARI	Professeur, Université Pierre et Marie Curie	Examineur
CHIHAB HANACHI	Professeur, Université de Toulouse 1	Examineur
LYNDA TAMINE	Professeur, Université de Toulouse 3	Directrice de thèse
WAHIBA BAHSOUN	Maître de Conférences, Université de Toulouse 3	Co-encadrante

École doctorale et spécialité :

MITT : Image, Information, Hypermedia

Unité de Recherche :

Institut de Recherche en Informatique de Toulouse (UMR 5505)

Directeur de Thèse :

Lynda TAMINE

Rapporteurs :

Eric GAUSSIER et Jian-Yun NIE

Définition et évaluation de modèles de recherche
d'information collaborative basés sur les
compétences de domaine et les rôles des
utilisateurs

Laure SOULIER

01/12/2014

Manuscrit soumis pour le diplôme de Docteur ès Sciences
Doctorante : Laure Soulier
Directrice de thèse : Professeur Lynda Tamine, Université de Toulouse 3 Paul
Sabatier
Co-encadrante : Maître de conférence Wahiba Bahsoun, Université de Toulouse
3 Paul Sabatier

© 2014 Laure Soulier

v2.2 2014-12-03

Commentaires, corrections, et autres remarques sont les bienvenus à :
{soulier,tamine,wbahsoun}@irit.fr

Institut de Recherche en Informatique de Toulouse, UMR 5505 CNRS,
Université Toulouse 3 Paul Sabatier,
118 route de Narbonne,
F-31062 Toulouse CEDEX 9

Remerciements

“L’ambition n’est pas signe d’orgueil, mais d’avenir !” Sans cette phrase qu’une amie m’a répétée à maintes reprises, je n’aurais jamais osé franchir le pas de cette porte qui ouvre un long chemin tumultueux, mais qui m’a appris tellement et m’a fait découvrir tellement de choses...

Je souhaite avant tout exprimer mes sincères remerciements à *Lynda Tamine-Lechani*, ma directrice de thèse, pour sa confiance et ses nombreux encouragements. Son soutien indéfectible, sa ténacité sans faille, ses conseils avisés m’ont été d’une grande aide tout au long de la thèse. Je ne la remercierai jamais assez pour tout le travail engagé, les longues heures de discussion et, encore aujourd’hui, son inquiétude envers mon devenir professionnel. Telle une seconde maman, elle a su faire émerger le meilleur de moi-même avec douceur et fermeté. Je la remercie également de m’avoir accordé sa confiance lors des enseignements que j’ai réalisés. Sa présence durant ces trois années a été source d’une grande motivation. Je souhaite exprimer ma sincère reconnaissance et mes remerciements à ma co-encadrante, *Wahiba Bahsoun*, enseignante depuis la troisième année de licence, qui m’a fait confiance dès le master et m’a guidé vers le chemin de la thèse. Je la remercie également pour ses relectures et sa bienveillance.

Je souhaite également remercier les rapporteurs de ma thèse, *Professeur Eric Gaussier (Université de Grenoble 1)* et *Jian-Yun Nie (Université de Montréal)*, pour avoir accordé du temps à une lecture attentive et détaillée de mon manuscrit ainsi que pour leurs remarques encourageantes et constructives. Je tiens également à remercier les examinateurs, *Professeur Florence Sedes*, *Professeur Maarten de Rijke*, *Professeur Patrick Gallinari* et *Professeur Chihab Hanachi*, d’avoir accepté de participer à mon jury de thèse et échangé leurs points de vue.

Je souhaite exprimer ma gratitude à *M. Claude Chrisment*, puis *Mme Josiane Mothe*, pour m’avoir accueilli au sein de l’équipe SIG. Je remercie l’ensemble des membres de l’équipe SIG pour leur gentillesse et les différents échanges que j’ai pu avoir avec chacun d’entre eux. Des remerciements tout particuliers à *Karen*, *Franck*, *Riad*, *Bernard* et *Yoann* de m’avoir accordé leur confiance en me laissant participer à leurs enseignements. Merci également à l’ensemble des doctorants, les échanges professionnels ou moins professionnels permettent d’égayer la vie d’un doctorant : *Madalina*, ma confidente ; *Dana*, mon tandem estival de rédaction ; *Biiiiiiieeeeeel :-)* , *Eya*, la chanteuse à toute heure ; *Adrian*, “give me five” ; *Faiza*, *Anthony*, *Rafik*, *Firas*, *Mohammed*, *Ismail*, *Arlind*, *Faten*, *Cyril*, et al.. *Lynda Said L’ hadj*, l’auto-motivation est un lien très fort entre nous, mais je crois qu’il en existe un encore plus grand : l’amitié...

Special thanks to Chirag Shah at SC&I (Rutgers) for accepting to collaborate during a three-month internship. Interesting perspectives and collaborations in CIR have been launched on the other side of the Atlantic Ocean. I learnt a lot and make wonderful friends : Chathra, Eric, Roberto, Vanessa, Kevin, Serife, ... Thank you to all of you for your support. Hope to see you somewhere, guys!

Il est une personne que je ne remercierai jamais assez... Dès le master, *Lamjed*, tu as été présent, attentif et toujours à l'écoute. Tes nombreux conseils, ton soutien à toute épreuve m'ont été d'une grande aide sur l'ensemble des travaux. Tes conseils m'ont permis d'acquérir rapidement des automatismes et des connaissances qui m'ont servi tout au long de ma thèse. C'est une grande leçon d'humilité et d'humanité d'avoir travaillé avec toi et de voir l'effet synergique de la collaboration, même indirecte. Sache que même si tu n'as pas contribué directement à mes travaux, il y a ta signature dans chacun d'entre eux. Merci...

Merci aux différents stagiaires avec qui j'ai échangé, tous ces échanges ont permis de faire avancer mes idées. La *dream team*, pour la partie développement du prototype et leurs encouragements; *Paul* pour son investissement et son amour pour Stromae! Je vous souhaite beaucoup de succès dans vos projets professionnels et personnels. Merci également à l'ensemble des participants impliqués dans les expérimentations utilisateurs. Votre collaboration a été précieuse.

En tant qu'ancienne représentante des doctorants à l'école doctorale MITT, je tiens à remercier l'ensemble du bureau pour ces nombreux échanges. Je remercie plus particulièrement Victor d'avoir pris le relais ces derniers mois.

Merci Xavier pour m'avoir accordé ta confiance pour les stages "Hippocampe".

La gaieté de mes amis, leur présence, leur soutien, leurs relectures ont été un moteur pour moi. Je ne vous remercierai jamais assez. *Steph*, ma géniale Washington girl qui m'a accompagnée pour ma première conférence, mais pour bien plus encore... *Coline*, ma kichou, merci pour ton éternelle fougue qui m'a permis d'avancer et de gravir des montagnes. *Marlène*, *Romain*, *Thomas*, *Laeitia*, *Franck*, *Clotilde*, *Maud*, *Marie*, *Fifi*, ... *AnneSo*, *Loïc*, *Jordan* et *Vincent*, promis les "back benchers", un jour, on retournera au fond de la classe...

Ma famille... Je crois que je n'aurais pas assez de mots pour vous remercier et vous dire ce que vous représentez pour moi. Vous avez cru en moi, vous êtes toujours là à mes côtés. Je n'aurais pu rêver d'une meilleure famille. Je suis fière de chacun d'entre vous. *Papa*, *Maman*, en un mot : MERCI. *Thomas*, fière de voir que tu deviens entrepreneur, propriétaire et mari de ma paxouille. *Manon*, ma demie, fière de toi tout court. Quoi que tu fasses, tu seras toujours le tiers d'Einstein. Je n'aurai pu rêver d'une meilleure "belle-doche", qui a pris le temps de relire ma thèse de fond en comble pour radier la moindre faute. *Alain*, merci d'avoir cru en moi, de m'avoir encouragé chaque jour. Et enfin, *mamie Jeanette*, je suis fière de te faire ce cadeau et de te voir assister à ma soutenance.

Last but not least. *Malek*, la thèse n'est pas un long fleuve tranquille, nous le savions, elle nous a permis de mieux nous connaître et de nous soutenir. Je suis convaincue qu'elle nous a rendus plus forts. Nous avons commencé le chemin ensemble et le terminons ensemble. L'avenir est à nous...

Abstract

The research topic of this document deals with a particular setting of information retrieval (IR), referred to as collaborative information retrieval (CIR), in which a set of multiple collaborators share the same information need. Collaboration is particularly used in case of complex tasks in which an individual user may have insufficient knowledge and may benefit from the expertise/knowledge or complementarity of other collaborators. This multi-user context rises several challenges in terms of search interfaces as well as ranking models, since new paradigms must be considered, namely division of labor, sharing of knowledge and awareness. These paradigms aim at avoiding redundancy between collaborators in order to reach a synergic effect within the collaboration process.

Several approaches have been proposed in the literature. First, search interfaces have been oriented towards a user mediation in order to support collaborators' actions through information storage or communication tools. Second, more close to our contributions, previous work focus on the information access issue by designing ranking models adapted to collaborative environments dealing with the challenges of (1) personalizing result set to collaborators, (2) favoring the sharing of knowledge, (3) dividing the labor among collaborators and/or (4) considering particular roles of collaborators within the information seeking process.

In this thesis, we focus, more particularly, on two main aspects of the collaboration :

- The expertise of collaborators by proposing retrieval models adapted to the domain expertise level of collaborators. The expertise levels might be vertical, in the case of domain expert and novice, or horizontal when collaborators have different subdomain expertise. We, therefore, propose two CIR models on two steps including a document relevance scoring with respect to each role and a document allocation to user roles through the Expectation–Maximization (EM) learning method applied on the document relevance scoring in order to assign documents to the most likely suited user.
- The complementarity of collaborators throughout the information seeking process by mining their roles on the assumptions that collaborators might be different and complementary in some skills. We propose two algorithms based either on predefined roles or latent roles which (1) learns about the roles of the collaborators using various search-related features for each individual involved in the search session, and (2) adapts the document ranking to the mined roles of collaborators.

Résumé

Cette thèse traite du domaine particulier de la recherche d'information (RI) appelé RI collaborative (RIC), qui considère un ensemble d'utilisateurs interagissant afin de satisfaire un besoin en information partagé. La collaboration est particulièrement utilisée dans le cas de tâches complexes ou exploratoires pour lesquelles les compétences d'un utilisateur unique peuvent être insuffisantes. La collaboration permet ainsi de rassembler les compétences complémentaires des utilisateurs afin de produire un résultat collectif qui dépasse l'agrégation des résultats individuels. Le contexte multi-utilisateurs ouvre de nombreuses perspectives en termes de support à la collaboration au travers la définition d'interfaces dédiées et de modèles d'ordonnancement basés sur les paradigmes de la collaboration. Ces derniers, dont principalement la division du travail, le partage des connaissances et la conscience de l'environnement collaboratif, doivent être pris en compte afin d'éviter la redondance entre les collaborateurs et de bénéficier de l'effet synergique de la collaboration.

Plusieurs approches ont été proposées dans la littérature. En premier lieu, les interfaces de recherche ont été orientées vers des médiations utilisateurs dans le but de proposer un support aux actions des collaborateurs au travers de moyens de communication et de stockage d'information. En deuxième lieu, plus proches de nos travaux, certains auteurs se sont intéressés à la problématique d'accès à l'information en définissant des modèles d'ordonnancement adaptés à l'environnement collaboratif afin de répondre aux enjeux suivants : (1) la personnalisation des ordonnancements aux collaborateurs, (2) l'amélioration du partage des connaissances, (3) la division du travail entre les collaborateurs et/ou (4) la considération des rôles des collaborateurs dans le processus de RI.

Dans ce manuscrit, nous nous intéressons particulièrement aux aspects suivants de la collaboration :

1. L'expertise de domaine des collaborateurs vis-à-vis du sujet du besoin en information partagé en proposant deux modèles d'ordonnancement collaboratifs basés (a) sur une distinction verticale des niveaux d'expertise avec les rôles d'expert et de novice de domaine et (b) une distinction horizontale de l'expertise où un groupe d'experts de sous-domaines collaborent pour répondre à un besoin en information à multiples facettes. Ces deux modèles sont composés de deux étapes principales qui estiment le score de pertinence de chaque document pour chacun des rôles puis alloue le document aux collaborateurs grâce à l'algorithme de la maximisation de l'espérance (EM).

L'objectif est d'assigner un document au collaborateur le plus à même de le juger en fonction de son niveau d'expertise.

2. La complémentarité des utilisateurs lors d'un processus de RIC au travers de méthodes d'apprentissage, en temps réel, des rôles des collaborateurs. Ceci a pour objectif de bénéficier des compétences pour lesquelles ils sont les plus efficaces et d'ordonnancer, ainsi, les documents en fonction des rôles des collaborateurs. Également, deux approches, supervisée et non supervisée, sont proposées permettant respectivement d'apprendre les rôles des collaborateurs en fonction de leurs actions. Ces approches sont basées respectivement sur (a) des rôles prédéfinis dans l'état de l'art et (b) des méta-rôles latents des collaborateurs non prédéfinis.

Publications

Article de journaux internationaux

1. **Laure Soulier**, Lynda Tamine, Wahiba Bahsoun. On Domain Expertise-based Roles in Collaborative Information Retrieval. In *Information Processing and Management (IP&M)*, Elsevier, Vol. 50 N. 5, p. 752–774, 2014.
2. **Laure Soulier**, Lamjed Ben Jabeur, Lynda Tamine, Wahiba Bahsoun. On Ranking Relevant Entities in Heterogeneous Networks Using a Language-Based Model. In *Journal of the American Society for Information Science and Technology (JASIST)*, Wiley, Vol. 64 N. 3, p. 500–515, 2013.

Articles de conférences internationales

1. **Laure Soulier**, Chirag Shah, Lynda Tamine. User-Driven System-Mediated Collaborative Information Retrieval. In *ACM SIGIR Special Interest Group on Information Retrieval (SIGIR 2014)*, p. 485–494, 2014.
2. **Laure Soulier**, Lynda Tamine, Wahiba Bahsoun. A Collaborative Document Ranking Model for a Multi-faceted Search. In *Springer Asia Information Retrieval Societies Conference (AIRS 2013)*, p. 109–120, 2013, Best paper award.
3. **Laure Soulier**, Lamjed Ben Jabeur, Lynda Tamine, and Wahiba Bahsoun. BibRank : a Language-Based Model for Co-Ranking Entities in Bibliographic Networks. In *ACM Joint Conference on Digital Libraries (JCDL 2012)*, p. 61–70, 2012.

Articles de conférences nationales

1. **Laure Soulier**, Lynda Tamine, Wahiba Bahsoun. Un modèle de recherche d'information collaborative basée sur l'expertise. In *Conférence en Recherche d'Information et Applications (CORIA 2014)*, p 35–50, 2014.
2. **Laure Soulier**, Lamjed Ben Jabeur, Lynda Tamine, and Wahiba Bahsoun. Modèle de langue pour l'ordonnement conjoint d'entités pertinentes dans un réseau d'informations hétérogènes. In *INformatique des Organisations et Systemes d'Information et de Décision (INFORSID 2012)*, 2012.

Table des matières

1	Introduction	13
1.1	De la recherche d'information individuelle à la recherche d'information collaborative	13
1.2	RI et collaboration	14
1.3	Recherche d'information collaborative : aperçu large du domaine	18
1.3.1	Fondements et paradigmes	18
1.3.2	Enjeux de la recherche d'information collaborative	19
1.4	Problématiques de la thèse	21
1.5	Contributions	21
1.6	Organisation de la thèse	23
I	Synthèse des travaux de l'état de l'art	27
2	Recherche d'information individuelle	31
2.1	Concepts de base de la RI	32
2.1.1	Concepts	32
2.1.2	Modèles de RI	34
2.1.2.1	Les modèles booléens	35
2.1.2.2	Les modèles vectoriels	36
2.1.2.3	Les modèles probabilistes	37
2.1.3	Evaluation	40
2.1.3.1	Le protocole d'évaluation	40
2.1.3.2	Les mesures d'évaluation	42
2.2	Recherche d'information orientée utilisateur	44
2.2.1	La modélisation du profil utilisateur	47
2.2.1.1	Centres d'intérêts de l'utilisateur	47
2.2.1.2	Expertise de l'utilisateur	50
2.2.1.3	Rôle de l'utilisateur	51
2.2.2	L'exploitation du profil utilisateur	52
2.2.2.1	Intégration des centres d'intérêts de l'utilisateur	53
2.2.2.2	Intégration de l'expertise de l'utilisateur	54
2.2.2.3	Intégration du rôle de l'utilisateur	55

2.2.3	Synthèse des approches de la RI orientée utilisateurs . . .	55
3	Recherche d'information collaborative	57
3.1	Recherche d'information et collaboration	58
3.1.1	Notion de collaboration	58
3.1.2	Motivations	61
3.1.3	Paradigmes de la RIC	62
3.1.3.1	La division du travail	62
3.1.3.2	Le partage des connaissances	63
3.1.3.3	La conscience de l'environnement collaboratif . .	64
3.2	Concepts de base	64
3.2.1	Collaborateurs	64
3.2.1.1	Le groupe de collaboration	65
3.2.1.2	Le collaborateur	65
3.2.2	La tâche collaborative	66
3.2.3	Le processus de RIC	67
3.3	Les modèles génériques de la collaboration	68
3.4	Domaines d'application	72
3.4.1	Domaine médical	72
3.4.2	Domaine de la gestion d'information	73
3.4.3	Domaine académique	74
3.5	Approches de la RIC	74
3.5.1	Approches basées sur la médiation utilisateur	75
3.5.1.1	Caractéristiques des approches de médiation utilisateur	75
3.5.1.2	Exemples d'interfaces collaboratives	76
3.5.1.3	Synthèse des caractéristiques des interfaces présentées	78
3.5.2	Approches basées sur la médiation système	79
3.5.2.1	Enjeux des modèles de RIC	80
3.5.2.2	Modèles d'ordonnancement collaboratif basés sur les rôles symétriques	80
3.5.2.3	Modèles d'ordonnancement collaboratif basés sur les rôles asymétriques	83
3.5.2.4	Synthèse des modèles de RIC	86
3.6	Evaluation de la RIC	88
3.6.1	Problématique de l'évaluation en RIC	88
3.6.2	Principales approches d'évaluation	89
3.6.2.1	L'évaluation basée sur la simulation	90
3.6.2.2	L'évaluation basée sur les logs des utilisateurs .	91
3.6.2.3	L'évaluation basée sur les expérimentations utilisateurs	92
3.6.3	Mesures d'évaluation	92
3.6.3.1	Mesures d'évaluation orientées système	93
3.6.3.2	Mesures d'évaluation orientées utilisateur	95

II Modèles de recherche d'information collaborative basés sur l'expertise 101

4 Modèles de recherche d'information collaborative basés sur l'expertise des collaborateurs	105
4.1 Contexte et questions de recherche	106
4.2 Ordonnancement collaboratif des documents selon des rôles basés sur l'expertise	108
4.2.1 Scénario de recherche	108
4.2.2 Modèle d'ordonnancement collaboratif basé sur les rôles d'expert et de novice de domaine	110
4.2.2.1 Notations : modélisation du collaborateur	112
4.2.2.2 Calcul des scores de pertinence des documents basés sur les rôles	115
4.2.2.3 Apprentissage des ordonnancements de documents basé sur les rôles d'expert et de novice de domaine	118
4.2.3 Modèle d'ordonnancement collaboratif basé sur les rôles d'experts de sous-domaines	122
4.2.3.1 Notations : modélisation thématique des entités à partir du besoin à multiples facettes	123
4.2.3.2 Identification des facettes du besoin en information partagé	124
4.2.3.3 Calcul des scores de pertinence basé sur les expertises de sous-domaines	128
4.2.3.4 Apprentissage des ordonnancements de documents basé sur les expertises de sous-domaines	129
4.3 Expérimentations et résultats	133
4.3.1 Cadre expérimental	134
4.3.1.1 Données expérimentales	134
4.3.1.2 Métriques d'évaluation	141
4.3.1.3 Modèles de référence	142
4.3.2 Validation du modèle de RIC basé sur les rôles d'expert et de novice de domaine	143
4.3.2.1 Analyse de l'efficacité au niveau de la session de recherche collaborative	144
4.3.2.2 Analyse de l'efficacité au niveau des rôles des utilisateurs	148
4.3.3 Validation du modèle de RIC basé sur un groupe d'experts de sous-domaines	152
4.3.3.1 Paramétrage du modèle	152
4.3.3.2 Analyse de l'efficacité de la recherche	154

III Modèles de recherche d'information collaborative

basés sur une médiation hybride utilisateur-système 161

5	Modèle d'apprentissage supervisé des rôles en contexte de col-laboration	165
5.1	Motivations et questions de recherche	166
5.1.1	Contexte	166
5.1.2	Analyse préliminaire	167
5.1.2.1	Méthodologie expérimentale	167
5.1.2.2	Notations : modélisation de la session collaborative	171
5.1.2.3	Etude du comportement des collaborateurs . . .	174
5.1.3	Contribution et questions de recherche	178
5.2	Notations : modélisation des patrons de rôles	180
5.3	Modèle hybride de RIC basé sur l'apprentissage des rôles des collaborateurs	181
5.3.1	Identification des différences de comportement entre col-laborateurs	182
5.3.2	Identification des rôles des collaborateurs	183
5.4	Expérimentation et résultats	187
5.4.1	Construction des jeux de données	187
5.4.1.1	Expérimentations utilisateurs	187
5.4.1.2	Collection de documents	188
5.4.1.3	Caractéristiques de comportement	188
5.4.1.4	Les rôles	189
5.4.1.5	Protocole et modèles de référence	191
5.4.1.6	Métriques d'évaluation	192
5.4.2	Résultats	193
5.4.2.1	Analyse des différences entre collaborateurs . . .	194
5.4.2.2	Analyse des changements de rôles détectés au cours de la session	195
5.4.2.3	Analyse de l'impact de l'apprentissage des rôles sur l'efficacité au niveau de la session	198
5.4.2.4	Analyse de l'impact de l'apprentissage des rôles sur l'efficacité au niveau de l'itération	201
6	Modèle de RIC non-supervisé basé sur les méta-rôles des col-laborateurs	203
6.1	Contexte et questions de recherche	203
6.2	Notations : modélisation des méta-rôles	205
6.3	Modèle non supervisé de collaboration basé sur les méta-rôles . .	207
6.3.1	Apprentissage des méta-rôles dans une session de RIC . .	208
6.3.1.1	Cadre d'apprentissage des méta-rôles des collaborateurs	209
6.3.1.2	Mesures de qualité des caractéristiques de comportement	210
6.3.1.3	Optimisation des méta-rôles	212

6.3.2	Ordonnancement collaboratif des documents basé sur les méta-rôles	216
6.4	Expérimentation et résultats	218
6.4.1	Construction du jeu de données	218
6.4.1.1	Expérimentations utilisateurs	218
6.4.1.2	Collection de documents	219
6.4.1.3	Scénarios et modèles de référence	219
6.4.1.4	Métriques d'évaluation	220
6.4.2	Résultats	221
6.4.2.1	Paramétrage du modèle de RIC basé sur les méta-rôles	221
6.4.2.2	Analyse de l'efficacité au niveau de la session . .	223
6.4.2.3	Analyse de l'efficacité au niveau de l'itération . .	226

IV Conclusion générale 231

7 Conclusion générale 233

7.1	Synthèse des contributions	233
7.2	Perspectives	235

Table des figures

1.1	L’aspect inter-disciplinaire de la RIC (Shah, 2013)	17
2.1	Processus en U de RI	33
2.2	Taxonomie des modèles de RI (Baeza-Yates and Ribeiro-Neto, 1999)	35
2.3	Représentation algébrique des documents et des requêtes dans l’espace des termes à deux dimensions	36
2.4	Protocole pour les campagnes d’évaluation officielles	41
2.5	Ensembles de documents utilisés pour l’évaluation d’un SRI	43
2.6	Domaines sous-jacents de la RI orientée utilisateur	45
2.7	Comparaison de l’espace documentaire des hypothèses de pertinence traditionnelle (Robertson et al., 1995) et personnalisée (Morris and Horvitz, 2007)	48
2.8	Modélisation des profils utilisateurs basée sur une approche connexionniste (Kim and Chan, 2003)	49
3.1	Les cinq niveaux de la collaboration selon Shah (2012)	60
3.2	Exemple d’une session de RIC impliquant trois utilisateurs	68
3.3	Étapes du modèle ISP (Kuhlthau, 1991)	69
3.4	Étapes fusionnées des deux modèles de RIC proposés par Gray (1989) et Marchionini (1995)	69
3.5	Modèle de RI sociale basé sur les interactions entre collaborateurs (Evans and Chi, 2010)	70
3.6	Modèle du comportement de l’utilisateur en contexte de RIC (Karunakaran et al., 2013)	71
3.7	Plugin <i>Coagmento</i> (Shah and González-Ibáñez, 2011a)	76
3.8	Interface <i>CoFox</i> (Rodriguez Perez et al., 2011)	77
3.9	Système <i>Fischlar-DiamondTouch</i> (Smeaton et al., 2006)	78
3.10	Système de collaboration reposant sur une table de travail interactive, respectivement TeamSearch (Morris et al., 2006) et WeSearch (Morris et al., 2010)	78
3.11	Principe de l’algorithme basé sur les rôles de <i>rassembleur</i> et de <i>surveilleur</i> (Shah et al., 2010)	85

3.12	Exemple de fichier log fourni par le participant RmitMG6 pour le sujet TREC 326i	90
3.13	Simulation d'une session collaborative à partir de logs de recherche d'une session interactive individuelle	91
3.14	Mesures d'évaluation pour la RIC (Shah, 2012)	93
3.15	Questionnaire du formulaire "TLX instrument"	96
3.16	Catégorisation des messages entre collaborateurs (González-Ibáñez et al., 2013)	97
4.1	Exemple d'un scénario de recherche impliquant deux collaborateurs	109
4.2	Aperçu général de notre modèle de RIC basé sur les rôles d'expert et de novice de domaine	112
4.3	Aperçu général de notre modèle de RIC basé sur les rôles d'experts de sous-domaines	122
4.4	Simulation d'une session de RIC à partir des fichiers log fournis par le participant rmitMG6 du sujet TREC 326i	140
4.5	Analyse de la mesure de précision $P@30$ au niveau des itérations pour les groupes de collaboration construits à partir d'une méthode exhaustive	145
4.6	Analyse de la mesure de précision $P@30$ au niveau des itérations pour les groupes de collaboration construits à partir d'une méthode sélective	147
4.7	Analyse de l'évolution du niveau d'expertise des novices au cours de la session de recherche	151
4.8	Paramétrage du facteur de diversité	153
4.9	Paramétrage du nombre de thèmes issus de la collection à partir du LDA	154
4.10	Paramétrage du nombre de facettes des requêtes collaboratives et de la fonction d'ordonnancement basée sur le LDA	154
4.11	Analyse de la micro-précision au niveau des itérations	156
4.12	Impact de la taille des groupes de collaboration sur l'efficacité	156
5.1	Barre d'outils du système de collaboration	168
5.2	Barre latérale du système de collaboration	168
5.3	Difficulté de la tâche exploratoire	169
5.4	Analyse des corrélations des caractéristiques de comportement des collaborateurs	175
5.5	Analyse quantitative de la difficulté des tâches	177
5.6	Contextualisation de notre approche par rapport à la notion de session hybride vis-à-vis de la médiation	179
5.7	Etape du modèle d'apprentissage des rôles	182
5.8	Comparaison de l'efficacité sur l'ensemble de la session de recherche.	201
6.1	Etapes du modèle collaboratif non supervisé basé sur les méta-rôles des collaborateurs	208
6.2	Méthodologie d'apprentissage des méta-rôles	209

6.3	Protocole d'évaluation basé sur la méthode de validation croisée	222
6.4	Paramétrage pour le scénario <i>MineRank</i> (q) selon la valeur γ	222
6.5	Paramétrage du scénario <i>MineRank</i> (t)	223
6.6	Analyse temporelle des caractéristiques de comportement impliquées dans l'étape d'ordonnancement collaboratif sur l'ensemble des sessions de recherche	225
6.7	Comparaison de l'efficacité au niveau des itérations sur l'ensemble des sessions	227

Liste des tableaux

1.1	Statistiques sur la fréquence de la collaboration dans une tâche de recherche d'information (Morris, 2013)	15
1.2	Comparaison des quatre principaux domaines de RI orientée utilisateur	16
2.1	Taxonomies des rôles en RIS et RIC	52
2.2	Synthèse des travaux de la RI orientée utilisateur	56
3.1	Statistiques sur la fréquence de collaboration (Morris, 2013) . . .	67
3.2	Synthèse des approches basées sur la médiation utilisateur	79
3.3	Synthèse des caractéristiques des modèles de RIC	87
3.4	Synthèse des caractéristiques des différents domaines de RI . . .	99
4.1	Notations utilisées pour l'algorithme d'allocation à la paire d'expert et de novice de domaine par méthode d'apprentissage	119
4.2	Notations utilisées dans l'algorithme d'allocation des documents aux experts de sous-domaines	129
4.3	Répartition des groupes d'experts selon leur effectif	136
4.4	Comparaison de l'efficacité de notre modèle collaboratif au niveau des sessions pour les groupes créés de façon exhaustive - % Tx : Taux d'accroissement de notre modèle. Test de significativité de student * : $0.01 < t < 0.05$; ** : $0.001 < t < 0.01$; *** : $t < 0.001$	144
4.5	Comparaison de l'efficacité de notre modèle collaboratif au niveau des sessions pour les groupes créés de façon sélective - % Tx : Taux d'accroissement de notre modèle. Test de significativité de student * : $0.01 < t < 0.05$; ** : $0.001 < t < 0.01$; *** : $t < 0.001$	146
4.6	Comparaison de l'efficacité de notre modèle collaboratif au niveau des rôles - % Tx : Taux d'accroissement de notre modèle. Test de significativité de student * : $0,01 < t < 0,05$; ** : $0,001 < t < 0,01$; *** : $t < 0,001$	149
4.7	Comparaison des moyennes de l'efficacité de notre modèle par rapport aux rôles des collaborateurs - % Tx : Taux d'accroissement de notre modèle. Test de significativité de student * : $0.01 < t < 0.05$; ** : $0.001 < t < 0.01$; *** : $t < 0.001$	150
4.8	Répartition des groupes d'experts pour la validation croisée . . .	152

4.9	Comparaison de l'efficacité de notre modèle - % Tx : Taux d'accroissement de notre modèle. Test de significativité de student * : $0.01 < t < 0.05$; ** : $0.001 < t < 0.01$; *** : $t < 0.001$	155
5.1	Catégorisation des réponses des participants à propos de la nature de la difficulté de la tâche	170
5.2	Statistique des expérimentations utilisateurs	171
5.3	Caractéristiques de comportement utilisées pour l'analyse préliminaire des comportements	172
5.4	Données statistiques des expérimentations utilisateurs	187
5.5	Caractéristiques de comportement pour modéliser la session . . .	189
5.6	Différences de comportement entre participants. $ Diff $: différence absolue. p -value ($p < 0.05$) : nombre de groupes avec des différences significatives entre les participants : * : au moins un groupe. ** : au moins 50% des groupes. *** : au moins 75% des groupes. **** : tous les groupes	194
5.7	Analyse de l'évolution des rôles détectés lors des sessions de recherche	197
5.8	Impact de la fenêtre temporelle d'apprentissage sur l'efficacité des sessions	198
5.9	Comparaison de l'impact de notre algorithme d'apprentissage des rôles sur l'efficacité d'une tâche de RIC en fonction des mesures orientées session. %Tx : taux d'amélioration de <i>RB-RIC</i> . Significativité du test de Student * : $0.01 < t \leq 0.05$; ** : $0.001 < t \leq 0.01$; *** : $t \leq 0.001$	199
5.10	Comparaison de l'impact de notre algorithme d'apprentissage des rôles sur l'efficacité d'une tâche de RIC en fonction des mesures orientées itération. %Tx : taux d'amélioration de <i>RB-RIC</i> . Significativité du test de Student * : $0.01 < t \leq 0.05$; ** : $0.001 < t \leq 0.01$; *** : $t \leq 0.001$	200
6.1	Caractéristiques de comportement	205
6.2	Notations utilisées pour l'apprentissage des méta-rôles	213
6.3	Analyse de l'efficacité de notre modèle basé sur l'apprentissage des méta-rôles lors d'une tâche de RIC. %Tx : amélioration de notre modèle <i>MineRank(q)</i> . Significativité du test de student * : $0.01 < t \leq 0.05$; ** : $0.001 < t \leq 0.01$; *** : $t \leq 0.001$	224

Chapitre 1

Introduction

1.1 De la recherche d’information individuelle à la recherche d’information collaborative

La recherche d’information (RI) est un domaine qui a pour objectif d’acquérir, d’organiser, de stocker et de rechercher l’information (Baeza-Yates and Ribeiro-Neto, 1999). Dans la plupart des travaux, la recherche d’information est perçue comme un processus individuel (Robertson and Walker, 1994; Salton and McGill, 1983) où un utilisateur unique formule un besoin en information par le biais d’une requête afin d’identifier des documents pertinents en rapport à ce besoin. Ainsi, un des enjeux majeurs du processus de RI réside dans la formalisation de modèle d’appariement requête-document (Ponte and Croft, 1998; Robertson et al., 1995; Salton and McGill, 1983), permettant de prédire la pertinence de chaque document d’une collection.

Dans ce contexte, certaines contributions (Buckley et al., 1994; Rocchio, 1971) se sont intéressées à la difficulté d’un utilisateur à exprimer son besoin afin de transmettre son intention de recherche (Jansen et al., 2008). L’objectif sous-jacent est d’améliorer le calcul du score de pertinence des documents grâce à des méthodes d’expansion ou de reformulation de requêtes basées sur les jugements de pertinence –*relevance feedback*– effectués par l’utilisateur au cours de sa session de recherche. Cette nouvelle approche de la RI a ouvert de nombreuses perspectives, centralisant ainsi l’utilisateur au sein du processus de RI et faisant émerger un nouveau domaine, appelé RI orientée utilisateur. Deux principales approches d’intégration de la dimension utilisateur en RI ont été proposées :

- L’approche individuelle orientée vers une personnalisation des résultats par rapport au contexte de l’utilisateur, ses préférences et ses intérêts selon des facteurs géographiques, temporels et comportementaux. Différentes notions sont alors abordées dans le domaine de la RI personnalisée (RIP), telles que

la modélisation du profil utilisateur (Kraft et al., 2005; Wen et al., 2004) et son intégration dans le calcul de la pertinence d'un document en réponse à une requête (Gauch et al., 2003; Liu et al., 2004).

- L'approche collaborative en considérant l'utilisateur au sein d'un groupe ou d'une communauté. Le calcul de la pertinence peut dépendre alors d'une collaboration implicite ou explicite. Dans le cas d'une collaboration implicite, on dénombre deux approches principales : (a) le filtrage collaboratif d'information où la pertinence est, par exemple, estimée à partir des préférences des utilisateurs similaires (Ma et al., 2009; Resnick et al., 1994) et (b) la recherche d'information sociale qui tient compte des interactions entre les utilisateurs, permettant ainsi de définir, par exemple, le degré de pertinence d'un document en fonction de l'autorité de son émetteur (Pal and Counts, 2011; Amer-Yahia et al., 2007). Lorsque la collaboration est explicite, les enjeux de l'estimation de la pertinence diffèrent car ils doivent prendre en compte le besoin en information partagé par un ensemble de collaborateurs dont l'objectif est d'interagir explicitement afin de le résoudre (Foster, 2006; Twidale et al., 1995).

1.2 RI et collaboration

De nombreuses applications collaboratives existent sur le web, telles que Wikipedia¹, Google Drive², Stack Overflow³, Doodle⁴ ou Trello⁵, facilitant la coordination entre les utilisateurs afin de générer du contenu de façon collective, de résoudre des problèmes via les forums ou de planifier des événements ou des tâches. Dans le cadre plus particulier de la RI, Morris (2013) a mis en évidence l'émergence de la collaboration dans des tâches de recherche, comme illustrée dans le Tableau 1.1 qui met en évidence la croissance quotidienne de la collaboration, atteignant 11% en 2012, contre 0.9% en 2006. Bien qu'elle soit de nature implicite, la collaboration est également perceptible dans les systèmes de questions-réponses ou les réseaux sociaux où le besoin en information émane d'un utilisateur qui requiert l'expertise d'un autre utilisateur pour répondre à son propre besoin initial (Hansen and Järvelin, 2005). Egalement, les utilisateurs interrogent leur réseau social lors des sessions de recherche exploratoires afin d'obtenir une information personnalisée et plus contextualisée par rapport au besoin en information soumis à un moteur de recherche (Oeldorf-Hirsch et al., 2014).

Comme énoncé précédemment, la collaboration en RI peut être considérée selon (a) un axe implicite où les données générées par les actions des utilisateurs sont inférées comme facteurs de pertinence (Foley and Smeaton, 2009) ou (b) un axe

1. <http://www.wikipedia.org/>

2. <http://drive.google.com/>

3. <http://stackoverflow.com/>

4. <http://doodle.com/>

5. <https://trello.com/>

Fréquence de la collaboration	2006	2012
Journalière	0.9%	11.0%
Hebdomadaire	25.7%	38.5%
Mensuelle	48.6%	15.6%
Moins fréquence	24.8%	34.8%

TABLE 1.1: Statistiques sur la fréquence de la collaboration dans une tâche de recherche d’information (Morris, 2013)

explicite où les utilisateurs expriment leur besoin de collaborer pour résoudre un besoin en information partagé (Foster, 2006). Trois principaux domaines de la RI basés sur la collaboration se distinguent :

- Le **filtrage collaboratif d’information (FCI)** (Belkin and Croft, 1992), également appelé recommandation, a pour objectif de personnaliser la RI pour un besoin formulé par un utilisateur unique à partir de données générées implicitement par d’autres utilisateurs. L’intuition principale est de s’adapter aux intérêts et aux attentes de l’utilisateur en recommandant des documents visités par les utilisateurs qui ont un profil similaire (Herlocker et al., 1999; Li and Kim, 2003; Pessiot et al., 2007; Resnick et al., 1994). La collaboration entre les utilisateurs est alors implicite.
- La **RI sociale (RIS)** (Korfiatis et al., 2006; Amer-Yahia et al., 2007) repose sur l’analyse des réseaux sociaux, modélisant les interactions entre les utilisateurs et les indicateurs sociaux liés aux documents et utilisateurs, afin d’améliorer l’estimation de la pertinence des documents (Heymann et al., 2008). Tout comme la FCI, la RIS repose sur une collaboration implicite entre les acteurs impliqués dans le calcul de la pertinence. Différentes tâches sont investies dans ce domaine, telles que la recherche de personnes, comme les experts d’un sujet (Balog et al., 2009), la recherche sociale (Kirsch et al., 2006) ou encore la détection des opinions (Osman et al., 2010) ou des sentiments (Rafrafi et al., 2012).
- La **RI collaborative (RIC)** (Twidale et al., 1997; Joho et al., 2009) consiste en la résolution d’un besoin en information partagé par un ensemble d’utilisateurs où chacun d’eux participe activement à l’identification des documents pertinents en interagissant avec les membres du groupe. Les interactions entre collaborateurs permettent de structurer la collaboration (Kelly and Payne, 2013) ainsi que de construire une réponse collective au besoin en information partagé (Foley and Smeaton, 2009).

Le Tableau 1.2 présente une comparaison des approches orientées utilisateur reposant sur une approche individuelle (RIP) ou collaborative (FCI, RIS et RIC) selon les dimensions suivantes : (a) le nombre d’utilisateurs, (b) la personnalisation, (c) le type de collaboration, (d) la concurrence des actions des utilisateurs,

		RIP	FCI	RIS	RIC
Utilisateur	unique	■	□	□	□
	groupe	□	■	■	■
Personnalisation	absente	□	□	■	■
	présente	■	■	■	■
Collaboration	implicite	□	■	■	□
	explicite	□	□	□	■
Concurrence	synchrone	□	□	□	■
	asynchrone	□	■	■	■
Bénéfice	symétrique	□	□	□	■
	asymétrique	□	■	■	□
Communication	absente	■	■	■	□
	présente	□	□	□	■
Usage de l'information	échange d'information	□	□	□	■
	recherche d'information	■	■	■	■
	synthèse d'information	■	□	■	■
	<i>sensemaking</i>	□	□	□	■

TABLE 1.2: Comparaison des quatre principaux domaines de RI orientée utilisateur

(e) le type de bénéfice que retire chaque utilisateur impliqué dans le processus, (f) l'absence ou la présence de communication et (g) l'usage de l'information. Pour cette dernière caractéristique, nous distinguons l'échange, la recherche, la synthèse d'information ainsi que le *sensemaking*, qui consiste à construire une connaissance à partir de l'information (Paul and Reddy, 2010; Tao and Tom-bros, 2013). Nous détaillons dans la section 3.1.1 certaines de ces différentes dimensions selon l'angle de la collaboration.

Dans cette thèse, nous nous intéressons au domaine de la RIC qui est particulièrement bénéfique pour la résolution de tâches complexes ou exploratoires (Singer et al., 2012; Shah, 2012), telles que la résolution d'un besoin en information à multiples facettes nécessitant l'exploration d'un ensemble de sous-sujets ou des tâches plus techniques, comme la planification d'un voyage, demandant

l'accord de plusieurs personnes (Morris and Teevan, 2009). En effet, pour ces tâches particulières, les connaissances et compétences d'un utilisateur unique ne sont pas toujours suffisantes (Denning and Yaholkovsky, 2008; Twidale et al., 1997). Ainsi, la principale motivation liée à la collaboration est de rassembler des compétences et connaissances complémentaires sous-jacentes à différents utilisateurs afin d'obtenir un bénéfice pour chacune des parties –*mutual benefits*–, aussi bien au niveau individuel que collectif (Shah, 2012). En outre, la collaboration favorise l'effet synergique de par la cohésion des actions des utilisateurs garantissant que le résultat agrégé de la collaboration est supérieur à la somme des résultats individuels (Shah and González-Ibáñez, 2011b). Dans ce contexte, la RIC est un domaine avec de nombreuses perspectives propres à la définition d'algorithmes d'ordonnancement collaboratif (Foley and Smeaton, 2009; Pickens et al., 2008), à l'analyse du comportement des utilisateurs (González-Ibáñez et al., 2011), mais également la conception d'interfaces adaptées à la collaboration (Rodriguez Perez et al., 2011; Smeaton et al., 2006) en termes de support au lien social entre les collaborateurs et à la gestion de l'information pertinente entre les collaborateurs. La RIC, appelée également *Collaborative Information Seeking (CIS)* dans certains travaux (Capra et al., 2010; Foster, 2006; Golovchinsky et al., 2009b), est généralement perçue comme une science interdisciplinaire (Foster, 2006; Shah, 2013), comme illustré dans la Figure 1.1.

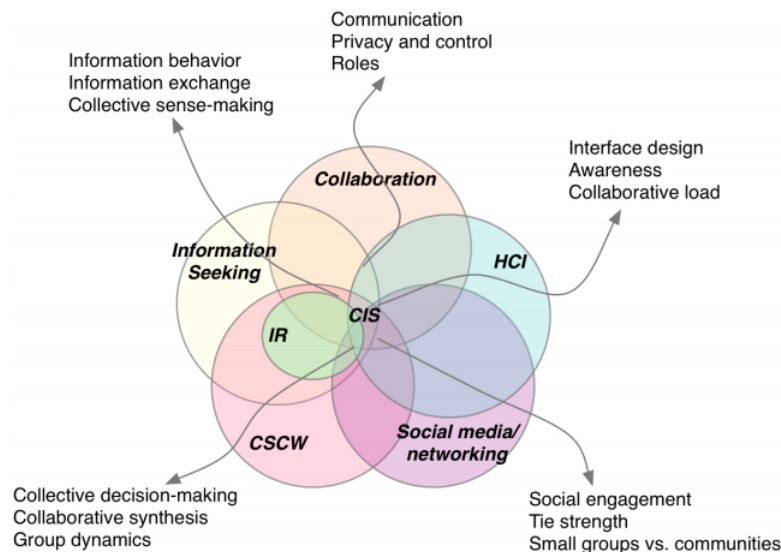


FIGURE 1.1: L'aspect inter-disciplinaire de la RIC (Shah, 2013)

Plus particulièrement, nos contributions s'intéressent à l'aspect RI du processus de collaboration, qui consiste en la définition de modèles d'ordonnancement permettant de prédire la pertinence des documents en rapport à un besoin en information partagé par un groupe de collaborateurs. De plus, dans la seconde

partie de notre contribution, nous introduisons la dimension du comportement des utilisateurs –*information behavior* dans la Figure 1.1– afin d’inférer en temps réel les rôles latents des collaborateurs lors d’une session de RIC.

1.3 Recherche d’information collaborative : aperçu large du domaine

1.3.1 Fondements et paradigmes

La RIC est définie comme une situation de recherche qui implique un ensemble d’utilisateurs qui se coordonnent afin de résoudre un besoin en information partagé (London, 1995). Le bénéfice de la collaboration est particulièrement avéré dans le cas de tâches complexes ou exploratoires afin de produire un effet synergique entre les utilisateurs qui dépasse la sommation des résultats individuels (Shah and González-Ibáñez, 2011b). La collaboration se manifeste selon différentes formes (Golovchinsky et al., 2009b) suivant la nature de la tâche et/ou la position géographique des collaborateurs, par exemple. De par son contexte multi-utilisateurs, le processus de RIC est caractérisé par des interactions hétérogènes (Evans and Chi, 2010) :

1. Les interactions collaborateurs-sources d’information qui permettent aux utilisateurs d’affiner et de donner du sens –*sensemaking*– (Paul and Reddy, 2010; Tao and Tombros, 2013) au besoin en information partagé.
2. Les interactions collaborateurs-collaborateurs qui permettent de structurer la collaboration ainsi que d’échanger et d’organiser les informations.

Afin de faciliter la collaboration dans un processus de recherche, deux principales approches sont utilisées (Joho et al., 2009). La première, de type médiation utilisateur, consiste à adapter les interfaces de recherche à un contexte multi-utilisateurs (Morris et al., 2006) en fournissant des outils dédiés, tels que les tables interactives (Smeaton et al., 2006), ou des logiciels proposant un espace de travail partagé avec un système de communication (Rodriguez Perez et al., 2011). La deuxième, plus proche de nos travaux, s’intéresse à la médiation système (Foley and Smeaton, 2009; Pickens et al., 2008) qui s’appuie sur les caractéristiques, les actions ou les rôles des utilisateurs afin de distribuer et d’ordonner les documents dans le contexte collaboratif pour améliorer les performances de recherche.

Dans ce contexte, la collaboration implique trois paradigmes principaux (Foley and Smeaton, 2010; Morris and Teevan, 2009) :

1. La conscience de l’environnement collaboratif –*awareness*– (Dourish and Bellotti, 1992; Gaver, 1991) qui permet aux collaborateurs de coordonner leurs actions. Ce paradigme est principalement garanti dans la définition d’interfaces de collaboration grâce à la création d’un espace de travail partagé, de

visualisation graphique des status et activités des membres du groupe de collaboration (Liechti and Sumi, 2002).

2. La division du travail –*division of labour*– (Foley and Smeaton, 2010) qui permet de garantir une séparation des tâches des collaborateurs et/ou des ensembles documentaires à explorer afin d’éviter la redondance entre leurs actions. Pour cela, les collaborateurs ont, par exemple, la possibilité de s’organiser au préalable par le biais de rôles ou d’une hiérarchie pour la coordination et l’optimisation de la session de recherche (Kelly and Payne, 2013).
3. Le partage des connaissances –*sharing of knowledge*– (Yao et al., 1999) qui assure la diffusion de l’information entre les collaborateurs. D’un point de vue de médiation utilisateur, il est mis en place sous la forme d’espaces de travail partagés ou de systèmes de communication (Poltrock et al., 2003) tandis que la médiation système utilise des modèles d’ordonnancement qui prennent en compte les jugements de pertinence des membres du groupe (Foley and Smeaton, 2009).

1.3.2 Enjeux de la recherche d’information collaborative

La RIC a été traitée dans beaucoup de travaux scientifiques (Hansen and Järvelin, 2005; London, 1995; Morris, 2013) depuis les deux dernières décennies, dont l’objectif est de répondre aux enjeux des deux types de médiation, utilisateur et système, soulignés par Joho et al. (2009).

Besoin en information partagé. En tant que première étape du processus de collaboration, la définition du besoin en information est primordiale dans la mesure où elle oriente les stratégies de recherche du groupe. Elle est, dans la majorité des cas, issue d’un processus humain. Cependant, le besoin en information n’est pas figé et peut évoluer au fur et à mesure de la session de recherche. En effet, les collaborateurs peuvent affiner leur besoin et donner du sens –*sensemaking*– (Paul and Reddy, 2010; Tao and Tombros, 2013) à leur recherche par le biais des interactions sociales et des connaissances enrichies à la lecture des documents. Un des principaux enjeux dans ce domaine est de fournir des outils, techniques ou méthodes permettant aux collaborateurs de faciliter cette démarche.

Satisfaction du bénéfice mutuel. Il est important qu’un processus de collaboration assure à la fois la satisfaction du groupe d’utilisateurs et celle de chacun des utilisateurs de façon intrinsèque. D’une part, une des solutions possibles en termes d’interface peut être de fournir un espace de travail qui permet la sauvegarde des documents au niveau collectif ainsi qu’au niveau individuel (Shah, 2012). Cette fonctionnalité permet ainsi à l’utilisateur de garder trace des documents qu’il a visités dans la session collaborative et qui l’intéressent à titre

individuel. D'autre part, une des forces des modèles d'ordonnancement collaboratif pourrait résider dans leur capacité à retourner des documents à chacun des membres du groupe en accord avec le besoin en information partagé, mais aussi prenant en compte ses caractéristiques intrinsèques, telles que les intérêts, les préférences ou les compétences des collaborateurs. L'analyse et la modélisation du comportement des utilisateurs est un élément de réponse à cet enjeu.

Coordination des collaborateurs. Une des difficultés de la RIC réside dans la coordination des actions des utilisateurs afin d'optimiser l'efficacité de la session de recherche. Trois principaux paradigmes sont identifiés comme leviers de coordination (Foley and Smeaton, 2010; Morris and Teevan, 2009) : (a) la division du travail qui permet d'éviter la redondance entre les actions des collaborateurs, (b) le partage des connaissances qui assure le flux d'information entre les collaborateurs et (c) la sensibilisation à l'environnement collaboratif –*awareness*–, qui favorise la prise de conscience des collaborateurs des actions des autres membres du groupe et des documents visités durant la session de recherche collaborative. Aussi bien au niveau interface qu'au niveau modèle d'ordonnancement, l'enjeu est de combiner ces trois paradigmes pour assurer une distribution optimale des actions des collaborateurs en fonction de la tâche de recherche. Il est important de noter que la collaboration induit automatiquement un coût cognitif supplémentaire d'organisation et de communication (Fidel et al., 2004). La qualité de coordination des collaborateurs est donc un enjeu primordial afin de leur permettre de compenser ce coût additionnel par un effet synergique en termes de résolution du besoin en information partagé.

Pertinence collective. L'estimation de la pertinence des documents est un élément clé pour la résolution du besoin en information partagé. En effet, en fonction de leurs connaissances, leurs compétences et leurs rôles dans la collaboration, les utilisateurs formulent des requêtes et visitent des documents différents, bien que toujours en rapport avec le besoin en information partagé. Ainsi, outre le facteur contenu, la pertinence collective d'un document en contexte collaboratif doit prendre en compte de nouveaux facteurs, tels que la division du travail et le partage des connaissances. En conséquent, la distribution des résultats de recherche serait optimisée ainsi que l'efficacité de la session collaborative. De plus, la configuration interactive de la session de recherche imposée par le contexte collaboratif nécessite la prise en compte des données générées par les utilisateurs afin d'être ré-utilisées dans les modèles d'ordonnancement par réinjection de pertinence. A l'inverse de la recherche d'information individuelle où les jugements de pertinence peuvent également participer à la définition de la pertinence individuelle (Joachims, 2002), l'enjeu repose ici sur la façon de considérer ces jugements de pertinence dans l'ordonnancement tout en garantissant la prise en compte des facteurs collaboratifs liés à l'environnement de recherche.

Evaluation des modèles de RIC. A ce jour, la RIC ne dispose pas de cadre formel pour l'expérimentation. Aucune collection de test officielle, composée d'un ensemble de documents, de requêtes tests et de jugements de pertinence associés, n'a été constituée pour l'évaluation de cette tâche particulière. Les mesures d'évaluation classiques ne tiennent pas compte de la tâche de collaboration et les facteurs à analyser, e.g. l'efficacité ou encore l'effort cognitif, ne sont pas clairement établis (Shah, 2013). Ainsi, afin de valider les modèles de RIC, la formalisation d'un cadre expérimental et la constitution de collections de test formelles deviennent un enjeu important.

1.4 Problématiques de la thèse

Cette thèse se concentre sur la définition de modèles de recherche d'information collaborative adaptatifs permettant de coordonner les utilisateurs et d'estimer la pertinence collective des documents, tout en garantissant la satisfaction du bénéfice mutuel de chacun des collaborateurs. Pour cela, nous proposons de prendre en considération les caractéristiques des collaborateurs selon deux aspects : leur niveau d'expertise par rapport au sujet de la requête collaborative et leur rôle dans la collaboration. Par conséquent, nous nous intéressons à deux problématiques principales :

1. Comment bénéficier des différences dans les compétences de domaine des collaborateurs afin d'optimiser l'effet synergique de la collaboration ?
 - (a) Comment inférer l'expertise des collaborateurs à partir de leurs activités de coordination ?
 - (b) Comment intégrer l'expertise des collaborateurs dans le modèle d'ordonnancement collaboratif ?
2. Comment exploiter le rôle comme dimension fondamentale de la division du travail dans une session de RIC ?
 - (a) Comment définir puis construire le rôle des collaborateurs en temps réel dans une session de recherche interactive ?
 - (b) Comment intégrer les rôles afin d'optimiser l'efficacité de la session de RIC ?

1.5 Contributions

La principale contribution de cette thèse consiste en la définition de modèles de RIC adaptatifs aux compétences de domaine et de recherche des collaborateurs. Pour cela, nous adoptons deux principales approches reposant respectivement sur les facteurs d'expertise et de rôle. La première approche s'intéresse à construire le modèle des collaborateurs selon leurs niveaux d'expertise par rapport au besoin en information partagé afin d'intégrer ce facteur dans des

modèles de RIC. La deuxième approche se concentre quant à elle, sur la vision dynamique des rôles des collaborateurs au travers de l'évolution de leur comportement au cours de la session et propose d'adapter les modèles d'ordonnement à cette dynamique des actions des utilisateurs. Pour chacune de ces approches, nous détaillons les contributions proposées ci-dessous.

Modèles collaboratifs basés sur l'expertise des utilisateurs. A ce jour, aucun des travaux du domaine de la RIC n'a pris en compte le niveau d'expertise des utilisateurs comme facteur intrinsèque dans l'inférence de la pertinence des documents. Ainsi, nous proposons deux modèles d'ordonnement reposant sur des différences de niveaux d'expertise entre collaborateurs perçues selon deux dimensions :

1. Dans un premier modèle, nous considérons une distinction verticale où nous faisons l'hypothèse qu'un utilisateur a un niveau d'expertise par rapport au besoin en information partagé plus élevé que son collaborateur. Une paire de rôles est ainsi définie : le couple d'expert et de novice de domaine. L'intuition de cette contribution est qu'en plus d'assurer les paradigmes de la collaboration, un utilisateur doit pouvoir accéder à des documents qui correspondent à son niveau d'expertise qui, plus particulièrement pour les novices (Agichtein et al., 2006; Privault et al., 2010), s'affine au fur et à mesure des documents visités.
2. Le second modèle s'intéresse à une distinction horizontale des niveaux d'expertise des collaborateurs, faisant l'hypothèse que chaque utilisateur est expert d'un sous-domaine du besoin en information et contribue à la résolution du besoin selon son domaine de compétence. L'objectif est alors de distribuer les documents en fonction des facettes d'expertise des collaborateurs tout en garantissant les paradigmes de la collaboration, et notamment celui de la division du travail.

Ces deux modèles ont été évalués selon un cadre expérimental basé sur la simulation de la collaboration à partir de fichiers logs de la collection TREC Interactive. L'effet synergique de ces modèles basés sur la prise en compte de l'expertise des collaborateurs a été démontré.

Modèles collaboratifs hybrides basés sur les rôles des collaborateurs. Dans cet axe, nous nous intéressons à la notion de rôle des collaborateurs comme facteur de division du travail et à son impact sur l'efficacité d'une session de recherche collaborative. Pour ce faire, nous menons en premier lieu une analyse exploratoire effectuée sur des fichiers log d'expérimentations utilisateurs –*user studies*– menées au sein de notre laboratoire. L'objectif de cette analyse est de comparer (a) des sessions de recherche où les actions des collaborateurs sont organisées selon des rôles prédéfinis et (b) des sessions de recherche libres où les utilisateurs s'organisent au fur et à mesure de la session. Une des conclusions majeures de cette analyse est que l'assignation des rôles prédéfinis contraint les utilisateurs dans leurs actions et ne permet pas d'optimiser pleinement la complémentarité

de leurs compétences. Pour résoudre cet aspect, nous proposons de définir un nouveau cadre de RIC qui, à la place d'ordonner les documents et d'assigner les rôles aux collaborateurs sans aucune connaissance a priori de leurs compétences de recherche, considère le comportement des collaborateurs de façon dynamique pour apprendre leur rôle en temps réel et adapter les modèles d'ordonnement en conséquence. Ainsi, une session de recherche est perçue comme une session hybride des deux niveaux de médiation (système et utilisateur), laissant l'utilisateur libre de ses actions dans la collaboration tout en le guidant avec un modèle en adéquation selon ses compétences et son comportement de recherche. Dans cette optique, nous suggérons deux principales approches :

1. Considérer les rôles définis dans l'état de l'art pour réinjecter en temps réel les modèles d'ordonnement associés.
2. Construire un modèle de rôles latents des collaborateurs, appelés aussi méta-rôles, non prédéfinis et non labelisés, qui sont ensuite utilisés dans une fonction de prédiction de pertinence basée sur ces méta-rôles.

Ces deux contributions ont été évaluées expérimentalement sur des fichiers log de sessions de recherche collaborative réelles, non simulées, et ont montré l'impact de l'apprentissage dynamique des rôles sur l'efficacité de la collaboration par rapport à des sessions de recherche collaborative où le rôle est prédéfini.

1.6 Organisation de la thèse

Cette thèse est constituée d'un chapitre introductif ainsi que de quatre principales parties, dont la première présente la synthèse des travaux de l'état de l'art, les deux suivantes détaillent nos principales contributions et la dernière conclut le manuscrit et discute des perspectives de recherche. Nous présentons le contenu ci-après.

Le *chapitre 1* introduit la thèse. Il présente le contexte, les problématiques de recherche abordées et les contributions issues de nos travaux.

La *première partie* de cette thèse, intitulée *Synthèse des travaux de l'état de l'art* présente le contexte de nos travaux. Compte tenu du cadre de notre thèse, nous avons axé l'état de l'art sur la dimension utilisateur en contexte de RI. Elle englobe deux chapitres :

- Le *chapitre 2* introduit les notions de base de la RI ainsi que les différentes approches de la RI orientée utilisateur.
- Le *chapitre 3* présente les principales notions de la RIC, les principales approches de la littérature relatives à la RIC ainsi que les cadres d'évaluation adaptés au contexte collaboratif.

La *deuxième partie* de cette thèse, intitulée *Modèles de RIC basés sur l’expertise*, présente la première partie de nos contributions relatives à l’intégration de l’expertise des collaborateurs dans un processus de RIC. Elle englobe un unique chapitre (*chapitre 4*) présentant deux modèles d’ordonnancement :

- Un modèle collaboratif d’ordonnancement des documents prenant en compte une distinction verticale des niveaux d’expertise des utilisateurs. Plus particulièrement, nous considérons les rôles d’expert et de novice par rapport au besoin en information partagé qui reposent sur des hypothèses de nouveauté (White and Dumais, 2009) et de spécificité (Kim, 2006) des ordonnancements. Dans l’objectif d’ajouter de la sémantique aux caractéristiques des rôles, nous proposons de modéliser, à partir des jugements de pertinence des collaborateurs, un profil d’expertise basé sur une représentation multinomiale des termes. Par conséquent, les caractéristiques des rôles ainsi que le profil d’expertise sont intégrés dans le calcul des scores de pertinence des documents. Un modèle d’ordonnancement collaboratif basé sur l’algorithme de la maximisation de l’espérance –*Expectation Maximisation (EM)*– (Dempster et al., 1977) permet ensuite d’allouer chaque document au collaborateur le plus à même de le juger. Ce chapitre inclut également l’évaluation expérimentale menée pour valider le modèle puis discute des résultats obtenus.
- Un modèle collaboratif d’ordonnancement des documents tenant compte, à la différence du modèle précédent, d’une distinction horizontale des niveaux d’expertise des utilisateurs. Dans ce contexte, nous considérons un groupe d’experts de sous-domaines de la requête collaborative afin de résoudre un besoin en information à multiples facettes. Grâce à l’algorithme d’Allocation Latente de Dirichlet –*Latent Dirichlet Allocation (LDA)* (Blei et al., 2003), nous identifions les facettes de la requête. Par conséquent, les documents et experts sont modélisés au travers d’une représentation thématique où chaque élément correspond à une facette de la requête. Nous proposons ensuite d’estimer le score de pertinence d’un document vis-à-vis du besoin en information et d’un expert grâce à une combinaison de scores de similarité respectivement basés sur les termes et les thèmes. Enfin, les documents sont distribués au groupe d’experts selon le paradigme de la division du travail. Similairement au modèle basé sur les rôles d’expert et de novice, l’ordonnancement collaboratif est réalisé grâce à l’adaptation du modèle d’apprentissage de maximisation de l’espérance. Egalement, nous présentons et discutons les résultats de notre validation expérimentale.

La *troisième partie* de cette thèse, intitulée *Modèles de RIC basés sur une médiation hybride utilisateur-système*, présente la deuxième partie de nos travaux relative à l’application du principe de division du travail grâce aux rôles des collaborateurs afin d’optimiser l’efficacité d’une session de RIC. Elle est structurée en deux chapitres :

- Le *chapitre 5* pose les fondements liés à la modélisation et exploitation du

rôle des collaborateurs au sein d’une session de RIC. Pour cela, nous présentons une analyse exploratoire dont l’objectif est de mettre en exergue l’impact du rôle sur le comportement des collaborateurs grâce à une analyse des corrélations des caractéristiques de comportement. Cette analyse repose sur trois expérimentations utilisateurs que nous avons menées au sein du laboratoire, impliquant 75 groupes de collaboration répartis respectivement entre un scénario de recherche libre et deux scénarios de recherche contraints à des rôles prédéfinis. La principale conclusion, qui guide nos contributions par la suite, fait émerger la nécessité de considérer le rôle des collaborateurs de façon dynamique du fait de l’évolution des compétences de recherche utilisées par les collaborateurs au cours de la session.

A partir de ce constat, nous proposons un modèle d’apprentissage supervisé des rôles des collaborateurs en situation de RIC. Notre modèle permet d’analyser les actions des collaborateurs au cours de la session, modélisées par des caractéristiques de comportement (Agichtein et al., 2006). Il repose sur trois étapes principales :

1. *L’identification des différences de comportement entre les collaborateurs.* Pour cela, nous modélisons le comportement des utilisateurs vis-à-vis de chaque caractéristique sous la forme d’un vecteur temporel qui recense leurs actions sur l’ensemble de la session. Les différences significatives sont alors identifiées grâce au test de Kolmogorov-Smirnov.
2. *L’inférence dynamique de leur rôle parmi un ensemble de rôles prédéfinis et labellisés.* A cet effet, nous modélisons des patrons de rôles qui opposent les spécificités de chacun des rôles prédéfinis par des corrélations négatives. Un appariement approximatif est alors réalisé entre la matrice de corrélations des collaborateurs qui modélise la complémentarité de leur compétences et la matrice de corrélations du patron de la paire de rôles.
3. *L’application du modèle d’ordonnancement collaboratif associé à la paire de rôles détectée* afin d’exploiter les compétences des collaborateurs dans la résolution de la tâche collaborative.

Nous terminons ce chapitre par l’évaluation expérimentale menée sur des fichiers log de sessions de recherche collaborative réelles, appelées également *–user studies–*.

- Le *chapitre 6* propose une extension du modèle d’apprentissage supervisé, présenté dans le chapitre 5. Plutôt que d’inférer le rôle des collaborateurs à partir d’un ensemble de rôles prédéfinis, notre modèle repose sur un apprentissage non supervisé qui identifie les rôles latents des collaborateurs à partir de leurs actions tout au long de la session de recherche. Les rôles latents sont appelés méta-rôles et sont modélisés par une matrice de corrélations basée sur les caractéristiques de comportement. Cette matrice est construite sur la base des caractéristiques pour lesquelles les collaborateurs ont été identifiés comme significativement différents, grâce au test de Kolmogorov-Smirnov. Les méta-rôles sont ensuite réinjectés dans un modèle d’ordonnancement,

grâce à un classifieur non supervisé de régression logistique en considérant comme facteurs d'apprentissage les caractéristiques des utilisateurs qui constituent le méta-rôle des collaborateurs. Enfin, nous présentons l'évaluation expérimentale menée sur des fichiers log d'expérimentations utilisateurs *-user studies-*, ainsi que les résultats obtenus, permettant de valider le modèle.

La *quatrième partie*, intitulée *Conclusion générale*, est composée d'un unique chapitre, à savoir le *chapitre 7*, qui discute de l'impact de nos contributions. Il conclut cette thèse et présente nos perspectives de recherche.

Première partie

Synthèse des travaux de l'état de l'art

Introduction

La RI s'intéresse particulièrement aux techniques et aux modèles qui permettent d'identifier l'information pertinente en rapport à un besoin utilisateur. Généralement, la RI est perçue comme un processus individuel où le besoin en information émane d'un seul utilisateur (Salton and Buckley, 1988). Ainsi, le système de RI (SRI) a pour objectif de satisfaire ce besoin en information individuel (Wilson, 1981).

Dans ce contexte, et plus particulièrement avec l'émergence du web 2.0, des travaux se sont intéressés à la place de l'utilisateur dans un processus de RI (Allan et al., 2003; Amer-Yahia et al., 2007), soulignant la nécessité d'adapter les modèles d'estimation de la pertinence à la requête ainsi qu'à l'utilisateur. En effet, de nouveaux facteurs de pertinence, outre celui du sujet de la requête, ont été proposés, afin de prendre en compte les caractéristiques intrinsèques des utilisateurs ainsi que son contexte de recherche (Golder and Donath, 2004; Lieberman, 1997). La RI est alors appelée RI orientée utilisateur. Elle englobe également certains travaux (Foley and Smeaton, 2009; Amer-Yahia et al., 2007) qui ont mis en évidence le besoin de collaboration entre les utilisateurs afin de bénéficier de l'expérience d'autres individus. D'une part, la collaboration peut être sous-jacente et implicite dans le cas où le besoin en information est résolu par un unique individu. La pertinence des documents dépend alors de facteurs collaboratifs impliquant (a) les historiques de recherche d'autres utilisateurs similaires (filtrage collaboratif d'information) ou (b) les méta-données ou données générées par les utilisateurs au sein d'un groupe ou d'une communauté (RI sociale). D'autre part, certaines tâches de recherche, complexes ou exploratoires, permettent de faire émerger une collaboration explicite où les utilisateurs partagent le même besoin et interagissent afin de le résoudre (Hansen and Järvelin, 2005). Dans ce domaine particulier, la collaboration est principalement motivée par le fait que les compétences d'un seul utilisateur unique ne peuvent suffire (Denning and Yaholkovsky, 2008). La tâche de RI devient alors collaborative nécessitant aux modèles, aux techniques et aux interfaces d'être adaptés à ce contexte multi-utilisateurs (Joho et al., 2009). Ce domaine est appelé RI collaborative et représente le coeur de notre travail de recherche.

Cette partie du manuscrit introduit les notions abordées tout au long de cette thèse. Dans un premier temps, nous présentons la RI individuelle de façon géné-

rale (Chapitre 2) en incluant les concepts fondamentaux ainsi que les techniques de prise en compte de l'utilisateur et de son contexte dans l'estimation de la pertinence. Dans un deuxième temps, nous nous intéressons à la RIC (Chapitre 3) en présentant ses motivations et notions sous-jacentes. Nous nous intéressons par la suite aux approches et aux techniques d'évaluation adaptées au cadre collaboratif.

Chapitre 2

Recherche d'information individuelle

Introduction

L'objectif principal de la recherche d'information (RI) consiste à relier les documents d'une collection avec un besoin en information formulé par un utilisateur. Il incombe aux modèles de RI d'estimer une mesure statistique qui permet de quantifier la pertinence d'un document par rapport à ce besoin. Les premiers modèles proposés en recherche d'information estiment cette pertinence seulement en fonction du sujet de la requête (Robertson et al., 1995). Or, deux utilisateurs soumettant la même requête peuvent attendre deux résultats différents, du fait (a) de leurs caractéristiques intrinsèques, telles que leur âge ou leur niveau d'étude, et (b) de leur contexte de recherche qui peut différer selon la dimension spatiale ou temporelle, entre autres. Par exemple, la requête *java* formulée par un utilisateur informaticien traduit probablement un besoin en information relié au langage de programmation tandis que, pour un utilisateur qui souhaite planifier ses vacances, la requête fait référence à l'île. Il semble donc opportun de considérer la modélisation de l'utilisateur dans le processus de recherche d'information. Dans ce contexte, deux dimensions de l'utilisateur s'opposent suivant s'il est perçu comme un individu unique, comme dans le domaine de la RI personnalisée (RIP), ou comme un individu au sein d'un groupe ou d'une communauté, comme dans les domaines du filtrage collaboratif d'information (FCI), de la RI sociale (RIS) ou collaborative (RIC).

Dans ce chapitre, nous présentons tout d'abord les fondements de la RI dans la section 2.1. La section 2.2 introduit l'émergence de la RI orientée utilisateur ainsi que ses principales approches.

2.1 Concepts de base de la RI

L'objectif d'un système de RI (SRI) est de relier le besoin en information d'un utilisateur, modélisé par une requête, avec un ensemble de documents en estimant leur pertinence par rapport à la requête (Salton and McGill, 1983). Dans ce qui suit, nous introduisons, dans un premier temps, le processus de RI ainsi que les notions sous-jacentes. Dans un second temps, nous détaillons les principaux modèles de RI utilisés pour l'estimation de la pertinence des documents. Enfin, nous présentons les cadres d'évaluation de la RI mis en oeuvre pour mesurer l'efficacité d'un SRI.

2.1.1 Concepts

L'objectif prépondérant de la RI est de satisfaire un besoin en information. Il est donc important que la requête qui modélise ce besoin en quelques mots traduise l'intention de l'utilisateur dans sa tâche de recherche. De nombreux travaux (Broder, 2002; Jansen et al., 2008; Kang and Kim, 2003; Rose and Levinson, 2004) se sont intéressés à cet élément-clé de la RI en définissant des taxonomies pour caractériser les objectifs de la tâche d'information exprimée par la requête, dont celle de Broder (2002) qui les classe selon trois dimensions :

- Les *requêtes informationnelles* qui ont pour objectif de chercher de l'information en rapport à un sujet donné, sans aucun *a priori* sur la source d'information. Ces requêtes, dont voici quelques exemples “gaz à effet de serre”, “météo à Toulouse”, “recette gâteau”, sont les plus courantes.
- Les *requêtes navigationnelles* qui consistent à atteindre une page web particulière, connue par l'utilisateur. Un exemple de requête peut être “université Paul Sabatier” si l'on souhaite accéder à la page d'accueil de l'université.
- Les *requêtes transactionnelles* qui souhaitent réaliser des transactions ou bénéficier de services en ligne. “Avion pas cher Toulouse-Paris” est un exemple de requêtes transactionnelles.

Le processus de RI qui permet, à partir d'une requête, d'ordonnancer les documents est appelé “processus en U”. Il est décomposé en trois principales étapes, illustrées dans la Figure 2.1 et détaillées ci-dessous.

1. L'**indexation** a pour objectif d'extraire le contenu d'un document dans un index (Salton and McGill, 1983; Zobel et al., 1998) où chaque terme du document est analysé et transformé pour devenir un mot-clé ou descripteur. L'atout de cette phase est de permettre un accès rapide à l'information contenue dans les documents. Le besoin en information de l'utilisateur, modélisé par une requête, est soumis au même traitement afin d'assurer une correspondance entre les deux entités à comparer. Il existe trois méthodes d'indexation :
 - (a) L'*indexation manuelle* est réalisée par un expert qui identifie les termes les plus représentatifs du document. Bien que subjective, cette méthode

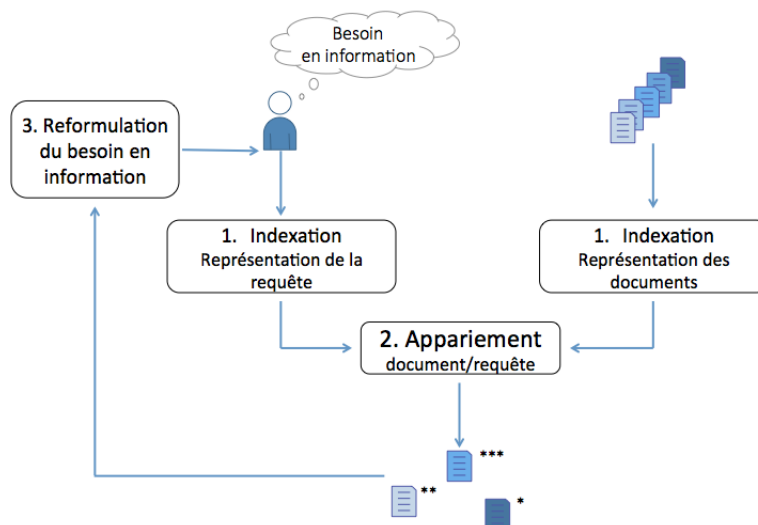


FIGURE 2.1: Processus en U de RI

présente l'avantage de fournir un vocabulaire contrôlé et assure une bonne qualité de représentation des documents.

- (b) L'*indexation automatique* repose sur une démarche algorithmique qui traite chaque terme selon un processus défini : extraction, suppression des mots vides, normalisation et pondération (Porter, 1997; Pirkola and Järvelin, 2001).
- (c) L'*indexation semi-automatique* est une combinaison des deux précédentes approches où le choix final des termes à indexer revient à l'expert.

A la fin de cette étape, les documents sont représentés dans des fichiers index qui stockent la cartographie des couples terme-document en y associant un poids. La formule de pondération la plus utilisée est celle basée sur la fréquence des termes dans les documents, appelée TF-IDF (Salton and Buckley, 1988). L'intuition de cette pondération est de favoriser les termes qui sont à la fois fréquents dans le document et peu fréquents dans la collection. Cette dernière condition est basée sur les propriétés de la loi de Zipf (1949) qui étudie la distribution des termes dans une collection de documents.

2. L'**appariement** correspond à l'étape qui met en relation la collection de documents, indexée au préalable, avec la requête, également pré-traitée, afin d'identifier les documents pertinents. Cette étape permet au SRI de retourner une liste de documents à l'utilisateur dont l'idéal est de faire correspondre deux types de pertinence (Cosijn and Ingwersen, 2000) :

- (a) La *pertinence système* qui est une mesure algorithmique basée sur le calcul de l'adéquation entre la représentation de la requête et celle de la collection de documents.
- (b) La *pertinence utilisateur* qui est la pertinence subjective que l'utilisateur aurait donné à chacun des documents.

Pour cette étape, le SRI dispose de modèles d'ordonnancement qui calculent la pertinence, appelée également score de similarité ou *Relevance Status Value (RSV)*, pour la paire document-requête.

3. La **reformulation du besoin en information** est l'étape qui permet de redéfinir le besoin de l'utilisateur au fur et à mesure de la session de recherche. Cette étape peut être effectuée :

- (a) Manuellement, dans le cas où l'utilisateur soumet lui-même une nouvelle requête.
- (b) De façon automatique, lorsque le SRI s'appuie sur les termes importants dans les documents les plus pertinents ou visités par l'utilisateur qui sont réutilisés :
 - i. Selon une boucle de rétropertinence *-relevance feedback-* qui considère les jugements de pertinence de l'utilisateur afin de construire une nouvelle requête traduisant plus finement le besoin en information de l'utilisateur (Rocchio, 1971).
 - ii. Selon une boucle de rétropertinence en aveugle *-pseudo-relevance feedback-* où le modèle utilise les n premiers documents retournés afin de construire la nouvelle requête (Cao et al., 2008; Xu et al., 2009).

2.1.2 Modèles de RI

L'étape d'appariement, décrite dans la sous-section 2.1.1, repose sur des modèles de RI dont l'objectif est d'identifier et d'ordonner les documents pertinents en rapport à un besoin en information. Une taxonomie des modèles a été présentée par Baeza-Yates and Ribeiro-Neto (1999) et présente quatre familles principales. Comme illustrée dans la Figure 2.2, les modèles reposent sur le texte des documents (*modèles de RI classiques* et *modèles basés sur le texte semi-structuré*), les liens entre documents (*modèles orientés web*) et les documents multimédia (*recherche d'images, de musiques, d'audio ou de vidéos*).

Compte tenu des concepts utilisés dans nos contributions, nous présentons dans cette sous-section les modèles appartenant à la catégories des *modèles de RI classiques* reposant respectivement sur la théorie des ensembles, les méthodes algébriques et les probabilités. Pour ce faire, nous considérons les notations suivantes :

- L'index \mathcal{I} est modélisé par le vecteur $\mathcal{I} = (t_1, \dots, t_v, \dots, t_V)$ où chaque élément t_v représente un terme de l'index \mathcal{I} et V correspond au nombre de termes dans l'index.

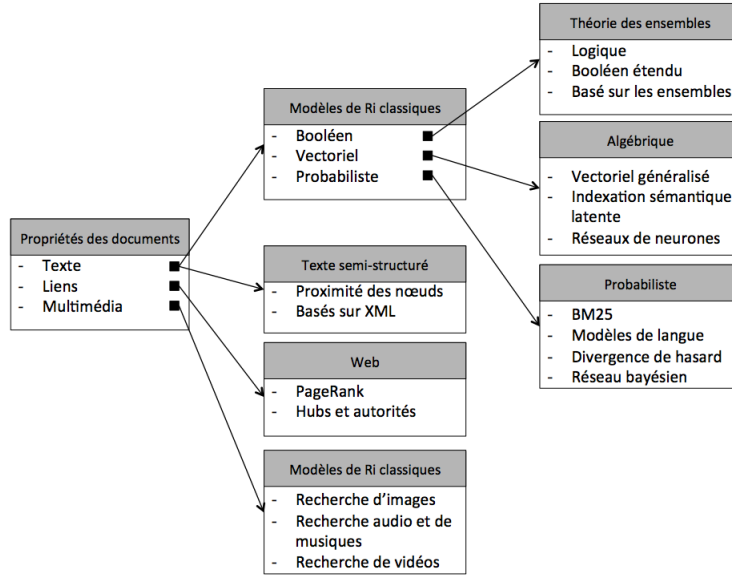


FIGURE 2.2: Taxonomie des modèles de RI (Baeza-Yates and Ribeiro-Neto, 1999)

- La collection de documents est notée $\mathcal{D} = \{d_1, \dots, d_i, \dots, d_N\}$ où N représente le nombre de documents dans la collection.
- L'ensemble de requêtes est noté $\mathcal{Q} = \{q_1, \dots, q_h, \dots, q_Z\}$ où Z représente le nombre de requêtes.
- Le document $d_i \in \mathcal{D}$ est modélisé vecteur de poids $d_i = (w_{i1}, \dots, w_{iv}, \dots, w_{iV})$ où chaque élément w_{iv} représente le poids du terme t_v pour le document d_i .
- La requête $q_h \in \mathcal{Q}$ est modélisée par un vecteur $q_h = (w_{h1}, \dots, w_{hv}, \dots, w_{hV})$ où chaque élément w_{hv} représente le poids du terme t_v pour la requête q_h .
- La fonction d'appariement est notée $RSV(d_i, q_h)$ et retourne le score de similarité du document d_i par rapport à la requête q_h .

Notons que, dans l'objectif de faciliter la lecture du manuscrit, nous adoptons la même notation pour faire référence à l'entité et sa représentation vectorielle.

2.1.2.1 Les modèles booléens

Le modèle logique, ou booléen, (Salton, 1971) est le modèle de base qui repose sur la théorie des ensembles et modélise les documents et les requêtes selon des ensembles de termes pondérés de façon binaire, correspondant à la présence ou l'absence du terme dans le document. Les termes au sein des documents et des requêtes sont reliés par des opérateurs logiques : OU (\vee), ET (\wedge) et NON (\neg).

Le score de similarité entre un document d_i et une requête q_h est inclus dans l'ensemble $\{0, 1\}$ et repose sur un appariement exact entre l'expression logique $\mathcal{L}(d_i)$ du document d_i et l'expression logique $\mathcal{L}(q_h)$ de la requête q_h , comme défini ci-après :

$$RSV(d_i, q_h) = \begin{cases} 1 & \text{si } \mathcal{L}(d_i) = \mathcal{L}(q_h) \\ 0 & \text{sinon} \end{cases} \quad (2.1)$$

L'avantage d'un tel modèle repose sur sa simplicité de mise en oeuvre et sa transparence à l'utilisateur dans la mesure où les documents sont retournés selon un appariement exact avec le besoin en information. Du fait de l'estimation binaire de la pertinence, ce modèle ne permet pas (a) d'identifier des documents partiellement pertinents qui pourraient répondre à certains aspects de la requête et (b) d'ordonner les documents selon leur degré de pertinence. Ces limites peuvent être corrigées par la considération non-binaire des poids des termes, utilisée dans le modèle booléen étendu (Salton and McGill, 1983) ou le modèle des ensembles flous (Ogawa et al., 1991).

2.1.2.2 Les modèles vectoriels

Le modèle vectoriel généralisé (Salton et al., 1975) repose sur la représentation algébrique des documents et des requêtes. Il a pour objectif de pallier aux faiblesses du modèle booléen classique qui repose sur un appariement exact en incluant la possibilité d'avoir un appariement approximatif avec différents degrés de pertinence entre les paires document-requête. Ce modèle représente les documents et requêtes dans l'espace vectoriel des termes de l'index, comme illustré dans la Figure 2.3 qui considère un index de deux termes.

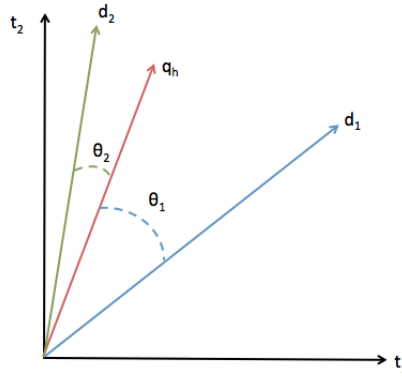


FIGURE 2.3: Représentation algébrique des documents et des requêtes dans l'espace des termes à deux dimensions

A partir de cette représentation algébrique, un score de similarité est estimé comme une mesure de proximité entre ces deux entités, correspondant à l'angle qui sépare les vecteurs requête et document. Il existe à ce jour plusieurs mesures dont les plus connues sont les suivantes :

1. Le produit scalaire où $RSV(d_i, q_h) = \vec{d}_i \cdot \vec{q}_h$,
2. La mesure de cosinus où $RSV(d_i, q_h) = \frac{\vec{d}_i \cdot \vec{q}_h}{\|\vec{d}_i\| \cdot \|\vec{q}_h\|}$ où $\|\vec{x}\|$ représente la norme euclidienne du vecteur \vec{x} ,
3. La mesure de Jaccard où $RSV(d_i, q_h) = \frac{|\vec{d}_i \cap \vec{q}_h|}{|\vec{d}_i \cup \vec{q}_h|}$ où $|\vec{d}_i \cap \vec{q}_h|$ correspond au nombre de termes présents à la fois dans la requête q_h et le document d_i tandis que $|\vec{d}_i \cup \vec{q}_h|$ représente le nombre de termes contenus dans la requête q_h ou le document d_i ,
4. La mesure de Dice où $RSV(d_i, q_h) = \frac{2 \cdot |\vec{d}_i \cap \vec{q}_h|}{|\vec{d}_i| + |\vec{q}_h|}$ où $|\vec{d}_i|$, respectivement $|\vec{q}_h|$, annote le nombre de termes dans le document d_i , respectivement la requête q_h .

En plus de bénéficier d'une mise en oeuvre facile, l'avantage de ce modèle est qu'il permet de considérer les documents partiellement pertinents et de les ordonner selon une fonction d'appariement approximatif. Cependant, il présente l'inconvénient de reposer sur l'hypothèse d'indépendance des termes –*bag of words*– alors que ce sont parfois les expressions ou les groupes de mots qui enrichissent la sémantique du document. Une des réponses à ce problème réside dans la considération des N-grammes (Song and Croft, 1999), permettant de regrouper des termes successifs qui peuvent avoir du sens ensemble.

Une autre solution, proposée par le modèle d'indexation sémantique latente (Furnas et al., 1988), consiste en l'utilisation des techniques d'analyse multidimensionnelle des termes, à savoir l'analyse en composante principale, afin de réduire les dimensions de représentation des documents et d'identifier les co-occurrences ou les corrélations entre les termes. L'appariement est alors effectué de façon approximative sur la base sémantique des documents.

2.1.2.3 Les modèles probabilistes

Nous décrivons dans ce qui suit, deux principales catégories de modèles probabilistes : les modèles probabilistes classiques et les modèles de langues.

Les modèles probabilistes classiques. Ces modèles (Robertson et al., 1995; Robertson and Walker, 1994) reposent sur la distribution de probabilité des termes pour identifier la similarité document-requête. Le principe de ce modèle (Robertson et al., 1995) est de favoriser les documents à la fois caractérisés par une forte probabilité d'être pertinents (événement P) et une faible probabilité

d'être non pertinents (évènement \bar{P}). Le score de pertinence d'un document d_i par rapport à la requête q_h est estimé comme suit :

$$RSV(d_i, q_h) = \frac{P(P|d_i)}{P(\bar{P}|d_i)} \quad (2.2)$$

où $P(P|d_i)$, respectivement $P(\bar{P}|d_i)$, représente la probabilité de pertinence, respectivement de non pertinence, par rapport à la requête q_h compte tenu du document d_i .

Cette fonction d'appariement peut être estimée ainsi :

$$RSV(d_i, q_h) = \prod_{t_v \in q_h} \frac{p_v(1-q_v)}{q_v(1-p_v)} \quad (2.3)$$

avec $p_v = \frac{r_v}{n}$ et $q_v = \frac{R-r_v}{N-n}$

où $t_v \in q_h$ correspond à l'ensemble des termes t_v de la requête q_h . p_v et q_v représentent respectivement la probabilité que le terme t_v apparaisse dans le document d_i sachant qu'il est pertinent, respectivement non pertinent, par rapport à la requête. Ces probabilités sont estimées par maximum de vraisemblance sur l'ensemble de la collection \mathcal{D} et dépendent du nombre total R de documents pertinents, dont r_v documents contenant t_v , ainsi que le nombre total N de documents dans la collection incluant n documents pertinents. Après développement et en ajoutant un coefficient de 0.5, afin d'éviter de diviser par 0, la formule finale, est la suivante :

$$RSV(d_i, q_h) = \prod_{t_v \in q_h} \frac{(r_v + 0.5)(N - n_v - R + r_v + 0.5)}{(n_v - r_v + 0.5)(R - r_v + 0.5)} \quad (2.4)$$

De nombreuses applications du modèle probabiliste ont été proposées dans la littérature, telles que le Okapi BM25 (Robertson and Walker, 1994) ou le modèle binaire BIR (Yu and Salton, 1976). Le modèle le plus utilisé est le modèle Okapi BM25 (Robertson and Walker, 1994). Les atouts majeurs de ce modèle consistent en la considération de la longueur des documents dans le calcul de la pertinence et de la fréquence des termes dans la collection, conformément à la loi de Zipf (1949). La fonction d'appariement est présentée dans l'équation ci-dessous :

$$RSV(d_i, q_h) = \sum_{t_v \in q_h} \frac{N - n_v + 0.5}{n_k + 0.5} \frac{f_{iv} \cdot (k_1 + 1)}{f_{iv} + k_1 \cdot (1 - b + b \cdot \frac{|d_i|}{avg_{dl}})} \quad (2.5)$$

où N représente la taille de la collection, n_v le nombre de documents qui contiennent le terme t_v . La fréquence du terme t_v dans le document d_i est notée f_{iv} . $|d_i|$ représente la longueur du document d_i tandis que la longueur moyenne des documents est notée avg_{dl} . Deux paramètres, respectivement k_1 et b sont utilisés et ont obtenu par expérimentation les valeurs optimales suivantes : $k_1 \in [1.2; 2.0]$ et $b = 0.75$.

Une autre extension dans cette catégorie de modèles repose sur des modèles informationnels (Clinchant and Gaussier, 2010; Na et al., 2008) qui analysent

la distribution de probabilité des termes. Pour cela, les auteurs se basent sur les heuristiques suivantes (Fang et al., 2004) : (a) plus les termes de la requête occurrent dans le document, plus le score du document est élevé, (b) l'occurrence marginale a plus d'importance lorsque la fréquence du terme dans le document est faible et (c) pour un même nombre d'occurrences, un document long reçoit un score plus faible. D'une façon générale, la fonction d'appariement document-requête est définie comme suit :

$$RSV(d_i, q_h) = \sum_{t_v \in q_h \cup d_i} -x_{t_v}^{q_h} \log \text{Prob}(X_{t_v} \geq t_{t_v}^{d_i} | \lambda_{t_v}) \quad (2.6)$$

où $t_v \in q_h \cup d_i$ correspond aux termes communs de la requête q_h et du document d_i . $x_{t_v}^{q_h}$ exprime le nombre d'occurrences du terme t_v dans la requête q_h et $t_{t_v}^{d_i}$ correspond à la version normalisée de la variable $x_{t_v}^{q_h}$. X_{t_v} est l'évènement lié à l'occurrence du terme t_v et λ_{t_v} est un paramètre lié à la distribution du terme t_v dans la collection. Dans ce contexte, Clinchant and Gaussier (2010) proposent d'utiliser deux distributions probabilistes afin d'estimer la probabilité $\text{Prob}(X_{t_v} \geq t_{t_v}^{d_i} | \lambda_{t_v})$:

— La distribution log-logistique, définie comme suit :

$$\text{Prob}(X_{t_v} \geq t_{t_v}^{d_i} | \lambda_{t_v}) = \frac{\lambda_{t_v}}{t_{t_v}^{d_i} + \lambda_{t_v}} \quad (2.7)$$

— La distribution de la loi de puissance lissée, définie ainsi :

$$\text{Prob}(X_{t_v} \geq t_{t_v}^{d_i} | \lambda_{t_v}) = \frac{\frac{t_{t_v}^{d_i}}{t_{t_v}^{d_i} + 1} - \lambda_{t_v}}{1 - \lambda_{t_v}} \quad (2.8)$$

Les modèles de langue. Le principe des modèles de langue (Ponte and Croft, 1998) repose sur le fait que la pertinence d'un document estime la similarité entre la requête q_h et le modèle de langue du document θ_{d_i} . Le score de similarité $RSV(d_i, q_h)$ est calculé comme suit :

$$RSV(d_i, q_h) = P(q_h | \theta_{d_i}) = \prod_{t_v \in q_h} P(t_v | \theta_{d_i}) \quad (2.9)$$

où $P(q_h | \theta_{d_i})$ représente la probabilité de la requête q_h sachant le modèle de langue θ_{d_i} du document d_i . Pour chaque terme t_v appartenant à la requête q_h , sa probabilité par rapport au modèle de langue θ_{d_i} du document d_i est notée $P(t_v | \theta_{d_i})$. Cette dernière probabilité s'appuie sur une estimation de la fréquence des termes de la requête q_h dans le document d_i mais est annulée pour les documents ne contenant pas tous les termes de la requête. Dans ce cas particulier, le score de similarité du document est nul alors que le document pourrait partiellement répondre au besoin en information formulé par la requête. Pour pallier à

cet inconvénient, des techniques de lissage ont été proposées (Jelinek and Mercer, 1980; MacKay and Peto, 1994). Ces dernières s'appuient sur un modèle de référence, en l'occurrence celui de la collection, pour estimer la pertinence d'un terme sur ce modèle de référence. Ces modèles de langue lissés ont montré leur efficacité dans de nombreuses applications en RI (Balog et al., 2009; Liang et al., 2012). Egalement, similairement à la prise en compte de la dépendance des termes au sein du modèle vectoriel, Maisonnasse et al. (2007) proposent de modéliser cette dépendance au travers d'un graphe afin de l'intégrer dans un modèle de langue.

2.1.3 Evaluation

L'évaluation d'un SRI permet de vérifier l'efficacité des modèles mis en oeuvre pour l'identification des documents pertinents. Dans cette section, nous présentons le cadre d'évaluation d'un SRI ainsi que les mesures d'évaluation sous-jacentes.

2.1.3.1 Le protocole d'évaluation

Dans l'objectif d'évaluer un modèle de RI, il est nécessaire de constituer des collections de test qui recoupent un ensemble de documents, des requêtes formulant des besoins en information et des jugements de pertinence associés qui recensent les documents pertinents pour chacune des requêtes. Cette approche d'évaluation correspond au paradigme de Cranfield (Cleverdon, 1997) qui a suscité le développement de nombreuses campagnes d'évaluation. Ces dernières présentent l'avantage de cibler une tâche particulière et d'évaluer l'efficacité des systèmes répondant à cette tâche. A l'initiative de cette démarche, le projet *Cranfield-1*, dirigé par Cleverdon, a constitué une collection de référence constituée de 18 000 documents et de 1 200 requêtes associées à un unique document pertinent. Par la suite, le projet *Cranfield-2* a établi des protocoles d'expérimentation basés sur l'hypothèse que les jugements de pertinence permettent de faire émerger la réponse globale au besoin en information sur un ensemble de participants et qu'une requête peut être associée à un ensemble de documents pertinents. Ainsi, les jugements de pertinence sont estimés selon une méthode de *pooling* où un ensemble d'individus, appelés juges, annotent manuellement les top documents retournés par chacun des participants de la campagne d'expérimentation. La technique de *pooling* utilisée lors des campagnes d'évaluation officielles est illustrée dans la Figure 2.4.

Les campagnes d'évaluation les plus connues sont :

1. La campagne TREC¹ –*Text REtrieval Conference*– est une des premières des campagnes qui regroupe à ce jour un large panel de tâches, telles que la

1. <http://trec.nist.gov>

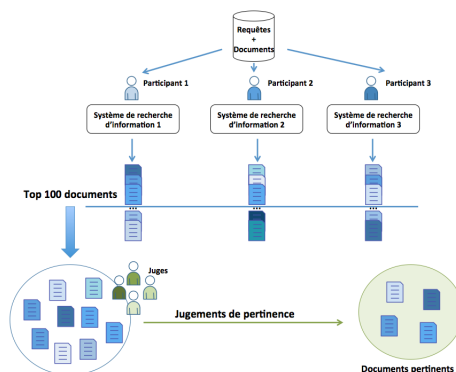


FIGURE 2.4: Protocole pour les campagnes d'évaluation officielles

recherche ad-hoc, ou également les tâches de recherche dans les microblogs ou celles orientées pour les systèmes de questions-réponses,

2. La campagne INEX² –*Initiative for the Evaluation of XML Retrieval*– oriente ses tâches de recherche vers des collections de documents structurés.
3. La campagne CLEF³ –*Conference and Labs of the Evaluation Forum*– propose des campagnes dans des langues différentes de l'anglais, traité majoritairement dans les campagnes TREC. En plus de proposer des tâches de recherche sur des documents, cette campagne fournit également des collections d'images associées à des annotations.

Dans ces campagnes, l'évaluation est plutôt centrée vers des tâches ad-hoc, ne considérant pas la dimension de l'utilisateur. L'émergence de la RI interactive et contextuelle (Dumais, 2009) a recentré l'utilisateur au sein du processus d'évaluation. De nouvelles tâches sont apparues, parmi lesquelles :

1. La tâche TREC Interactive⁴ qui consiste en la résolution d'un besoin complexe. Les participants doivent alors fournir les fichiers log qui recensent les interactions des utilisateurs (requêtes soumises, documents visités, ...).
2. La tâche TREC Contextual Suggestion⁵ qui consiste en une suggestion de documents à partir d'un ensemble de profils utilisateur et d'un contexte, traduisant respectivement les préférences et la localisation des utilisateurs.
3. La tâche TREC Session Search⁶ qui consiste en l'ordonnancement des documents vis-à-vis d'une requête particulière, soumise à un moment donné de la session, à partir de l'historique de recherche antérieur d'un utilisateur (requêtes reformulées antérieurement et leurs ordonnancements et jugements de pertinence associés).

2. <https://inex.mmci.uni-saarland.de>

3. <http://www.clef-initiative.eu>

4. <http://trec.nist.gov/data/interactive.html>

5. <https://sites.google.com/site/trecontext/>

6. <http://ir.cis.udel.edu/sessions/>

En dehors de ces campagnes d'évaluation, des protocoles d'évaluation ont été proposés afin d'intégrer la dimension de l'utilisateur. Nous dénombrons quatre types d'évaluation (Dumais, 2009; Kelly et al., 2009) :

1. Les expérimentations basées sur la simulation (Sieg et al., 2007) qui permettent de simuler le comportement des utilisateurs en construisant le scénario de recherche afin d'évaluer le modèle selon des cas d'étude *–use case–* bien particuliers.
2. Les expérimentations basées sur les fichiers logs d'expérimentations réelles *–log study–* (Anick, 2003) qui permettent d'évaluer un modèle de RI sur des données réelles. Ces expérimentations ne reposent pas sur des jugements explicites mais infèrent implicitement le comportement des utilisateurs à partir des fichiers logs.
3. Les expérimentations utilisateurs *–user study–* (Ringel et al., 2003; Yee et al., 2003) qui sont généralement effectuées en laboratoire où les participants sont confrontés au modèle de RI en situation réelle de recherche. Les tâches de recherche sont alors contrôlées et les résultats en terme d'efficacité sont à nuancer par rapport au facteur comportemental des individus.
4. Les expérimentations basées sur les études des journaux *–diary studies–* (Boudghaghen et al., 2010) qui reposent sur des sessions de recherche naturelles où les utilisateurs réels initient le besoin en information et réalise la session de manière libre et non contrainte.

2.1.3.2 Les mesures d'évaluation

Les mesures d'évaluation permettent d'estimer quantitativement l'efficacité d'un système. L'objectif est d'identifier, pour chaque requête la capacité du système à retourner des documents pertinents. La Figure 2.5 illustre les différents ensembles manipulés lors de l'évaluation d'un SRI, à savoir les ensembles des documents pertinents et des documents retournés par le système. Les documents pertinents non retournés par le système constituent l'ensemble de documents *silence* tandis que les documents non-pertinents retournés par le système génèrent du *bruit*. La principale difficulté d'un système de RI est de reposer sur un modèle qui retourne le maximum de documents pertinents (*minimiser le silence*) sans augmenter le nombre de documents non pertinents retournés (*minimiser le bruit*).

Chaque requête q_h est évaluée individuellement au moyen d'une mesure statistique estimée au rang r de la liste l_h retournée par le SRI. La mesure est ensuite agrégée sur l'ensemble des requêtes de la collection de test afin d'obtenir la mesure d'efficacité moyenne du système. Nous détaillons dans ce qui suit les principales mesures d'évaluation au rang r . Deux catégories de mesures sont identifiées dans la littérature.

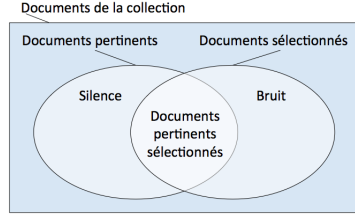


FIGURE 2.5: Ensembles de documents utilisés pour l'évaluation d'un SRI

Les mesures orientées rappel et précision. Elles évaluent l'efficacité des ordonnancements en se basant sur le nombre de documents pertinents retournés par le système. Nous distinguons trois principales mesures :

1. Le rappel $Rappel@r$ mesure la capacité d'un système à retourner tous les documents pertinents au rang r , et par conséquent, sa capacité à minimiser le *silence*, illustré dans la Figure 2.5. Pour une requête donnée, il est estimé par le ratio du nombre de documents pertinents retournés par le système par rapport au nombre de documents pertinents pour cette même requête. Ce ratio est ensuite agrégé sur l'ensemble des requêtes $q_h \in Q$:

$$Rappel@r = \frac{1}{|Q|} \sum_{q_h \in Q} Rappel(q_h)@r = \frac{1}{|Q|} \sum_{q_h \in Q} \frac{|Sel_{q_h}@r \cap Pert_{q_h}|}{|Pert_{q_h}|} \quad (2.10)$$

où $Sel_{q_h}@r$ regroupe l'ensemble des documents sélectionnés par le SRI pour la requête q_h au rang r . $Pert_{q_h}$ représente l'ensemble des documents pertinents pour la requête q_h .

2. La précision $Precision@r$ au rang r mesure la capacité d'un système à ne retourner que des documents pertinents, à savoir sa capacité à minimiser le *bruit*, illustré dans la Figure 2.5. Pour une requête donnée, elle est estimée par le ratio du nombre de documents pertinents retournés par le système par rapport au nombre de documents sélectionnés pour cette même requête. Ce ratio est ensuite agrégé sur l'ensemble des requêtes $q_h \in Q$:

$$Precision@r = \frac{1}{|Q|} \sum_{q_h \in Q} Precision(q_h)@r = \frac{1}{|Q|} \sum_{q_h \in Q} \frac{|Sel_{q_h}@r \cap Pert_{q_h}|}{|Sel_{q_h}@r|} \quad (2.11)$$

3. La F-mesure $F - mesure@r$ combine les mesures de rappel et de précision estimées au rang r afin de mesurer la capacité du système à retourner le maximum de documents pertinents, et seulement ces documents-ci, traduisant ainsi le double objectif d'un système de RI de minimiser le bruit et le silence. La F-mesure est estimée comme suit :

$$F - mesure@r = \frac{1}{|Q|} \sum_{q_h \in Q} 2 \cdot \frac{Rappel(q_h)@r * Precision(q_h)@r}{Rappel(q_h)@r + Precision(q_h)@r} \quad (2.12)$$

Les mesures orientées rang. Elles ajoutent un niveau d'analyse supplémentaire en mesurant la fiabilité des ordonnancements grâce au rang des documents pertinents. Trois mesures principales sont utilisées :

1. La précision moyenne $MAP@r$ –*Mean Average Precision*– correspond à la précision moyenne obtenue à chaque rang considéré par rapport aux r premiers résultats de la liste l_h retournée pour la requête q_h :

$$MAP@r = \frac{1}{|Q|} \sum_{q_h \in Q} \frac{1}{r} \sum_{R=1}^r Precision(q_h)@R \quad (2.13)$$

2. Le rang inverse moyen $MRR@r$ au rang r –*Mean Reciprocal Rank (MRR)*– favorise l'hypothèse que les documents pertinents doivent être retournés en premier dans la liste d'ordonnement par rapport aux r premiers documents retournés. Ainsi, cette métrique estime le rang moyen $Rang(l_h)@r$ du premier document pertinent dans les listes de résultats $l_h \in L$ de r documents retournés en réponse à la requête q_h :

$$MRR = \frac{1}{|Q|} \sum_{q_h \in Q} \sum_{l_h \in L} \frac{1}{Rang(l_h)@r} \quad (2.14)$$

3. Le NDCG au rang r , noté $NDCG@r$ –*Normalized Discounted Cumulative Gain*– évalue la capacité d'un système à retourner des documents pertinents par degré de pertinence (Järvelin and Kekäläinen, 2002). À l'inverse des mesures précédentes où les jugements de pertinence sont binaires (pertinents ou non), cette mesure repose sur des jugements de pertinence graduels. Elle peut être estimée à partir de la mesure de DCG –*Discounted Cumulative Gain*– appliquée à la liste l_h de résultats normalisée par la même mesure appliquée à l'ordonnement idéal des jugements de pertinence en fonction de leur degré, notée $IDCG(q_h)$:

$$NDCG@r = \frac{\sum_{q_h \in Q} DCG(l_h)@r}{\sum_{q_h \in Q} IDCG(l_h)@r} \quad (2.15)$$

$$DCG(l_h)@r = rel_1 + \sum_{R=1}^r @R \frac{rel_R}{\log_2 R}$$

avec rel_R qui correspond au jugement effectué par l'utilisateur au document situé au rang R de la liste L_h de résultats.

2.2 Recherche d'information orientée utilisateur

Dans la section précédente, nous avons abordé la RI selon le paradigme “*one size fits all*” qui considère que la même liste de documents est destinée à répondre à un même besoin en information formulé par deux utilisateurs différents. La fonction d'appariement d'un document d_i dépend seulement de la requête q_h et est notée $RSV(d_i, q_h)$. Au cours de ces dernières années, le besoin de considérer

certain aspects implicites d'une session de recherche, tels que les jugements de pertinence *-relevance feedback-* (Joachims, 2002; Rocchio, 1971), a émergé afin d'améliorer la performance des SRI. L'utilisateur est alors placé au centre du processus de recherche, faisant ainsi émerger la RI orientée utilisateur qui peut être analysée selon plusieurs domaines. Ces derniers ont été présentés dans la section 1.2 et illustrés dans la Figure 2.6 :

1. La RI personnalisée (RIP) dont l'objectif est d'adapter les ordonnancements aux caractéristiques de l'utilisateur, telles que ses préférences et intérêts (Agichtein et al., 2006; Sieg et al., 2004).
2. Le filtrage collaboratif d'information (FCI) qui consiste en la recommandation de documents à partir des actions d'utilisateurs qui partagent des intérêts similaires (Li and Kim, 2003; Mooney and Roy, 2000).
3. La RI sociale (RIS) qui s'appuie sur les interactions entre utilisateurs afin d'identifier les documents pertinents ou les acteurs importants (Amer-Yahia et al., 2007; Kirsch et al., 2006).
4. La RI collaborative (RIC) dont l'objectif est de distribuer les documents par rapport à un contexte multi-utilisateurs dont les membres sont rassemblés par la nécessité de résoudre ensemble un besoin en information partagé (Foster, 2006; Twidale et al., 1995). Cette branche de la RI fait l'objet d'une étude détaillée dans le Chapitre 3.

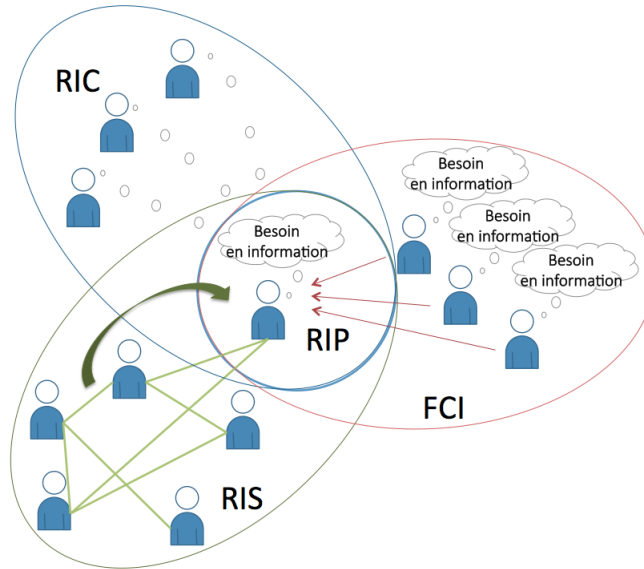


FIGURE 2.6: Domaines sous-jacents de la RI orientée utilisateur

Dans ce contexte, certains auteurs (Allan et al., 2003; Kirsch et al., 2006; Plickens et al., 2008) énoncent la nécessité de définir de nouveaux modèles de RI orientés utilisateur, permettant d'adapter la fonction d'estimation de la pertinence des documents à la dimension de l'utilisateur u_j , qu'elle soit considérée

dans un processus individuel ou collectif. Le score de pertinence peut être noté $RSV(d_i|q_h, u_j)$. Ainsi, deux enjeux majeurs se distinguent dans la RI orientée utilisateur :

1. La modélisation du profil utilisateur qui regroupe les problématiques d’acquisition et de représentation des caractéristiques de l’utilisateur. Pour l’acquisition des caractéristiques, deux catégories de données peuvent être exploitées :
 - (a) Des données explicites collectées directement de l’utilisateur au travers, par exemple, de formulaires (Hupfer and Detlor, 2006; Martin and Jose, 2004).
 - (b) Des données implicites dérivées des activités de l’utilisateur (Liu et al., 2004; Shen et al., 2005; Dumais et al., 2003), appelées également *relevance feedback*, au travers des pages visitées ou des requêtes soumises, par exemple.

La représentation de ces caractéristiques de l’utilisateur peut être matérialisée selon des dimensions variées, telles que ses centres d’intérêts, son niveau d’expertise ou son rôle, qui sont définies dans des taxonomies de modélisation de l’utilisateur (Göker and Myrhaug, 2002; Ingwersen and Järvelin, 2005; Tamine-Lechani et al., 2010).

2. L’exploitation du profil dans les modèles d’ordonnancement afin d’adapter les résultats au profil de l’utilisateur, en plus du besoin en information. Dans le cas d’un processus individuel personnalisé (RIP), l’exploitation se traduit par une personnalisation des ordonnancements (Teevan et al., 2005; Radinsky et al., 2013) tandis que lors d’un contexte collaboratif, la pertinence est estimée implicitement en fonction des autres utilisateurs dans le domaine du FCI et de la RIS (Ben Jabeur et al., 2010; Larson et al., 2011), ou de façon explicite pour le domaine de la RIC (Morris et al., 2008; Pickens et al., 2008).

Compte tenu de la thématique de cette thèse axée sur la collaboration multi-utilisateurs qui nécessite la prise en compte du comportement des collaborateurs et de leurs interactions durant le processus de recherche (Agichtein et al., 2006; González-Ibáñez et al., 2013), nous nous intéressons dans cette section plus particulièrement à l’analyse des données implicites qui permettent de comprendre et d’exploiter les actions des utilisateurs en tenant compte de différents facteurs. Plus particulièrement, nous nous intéressons aux facteurs liés aux intérêts et aux préférences des utilisateurs, à leur niveau d’expertise ainsi qu’à leur rôle qui sont essentiels au processus de collaboration dans la mesure où ils permettent d’assurer une complémentarité des compétences entre collaborateurs (Sonnenwald et al., 2004) ainsi que de structurer le processus (Kelly and Payne, 2013) afin d’éviter la redondance entre les actions. Nous présentons, dans ce qui suit les deux enjeux de la RI orientée utilisateurs en fonction de ces trois facteurs de centres d’intérêts, d’expertise et de rôle.

2.2.1 La modélisation du profil utilisateur

Nous présentons dans cette section les différents travaux liés à la modélisation du profil utilisateur selon les facteurs de centres d'intérêts, d'expertise et de rôle.

2.2.1.1 Centres d'intérêts de l'utilisateur

D'une manière générale, les centres d'intérêts font partie de la dimension cognitive de l'aspect utilisateur (Oard, 2001; Tamine-Lechani et al., 2010). Dans ses travaux, Park (1994) souligne l'importance des centres d'intérêts des utilisateurs dans la résolution des besoins en information. Plus particulièrement lorsqu'ils sont ambigus, la prise en compte des intérêts de l'utilisateur ayant formulé ce besoin en information peut permettre d'identifier quel aspect de la requête l'intéresse.

L'hypothèse principale des modélisations basées sur les centres d'intérêts a été formulée par Oard (2001), considérant que les interactions d'un utilisateur avec le système dénotent des jugements d'intérêt de l'utilisateur vis-à-vis du besoin en information. On dénombre trois principales approches de modélisation des centres d'intérêts en RI :

1. *Les approches basées sur les termes* (Chen et al., 1998; Lieberman, 1997; Sakagami and Kamba, 1997; Tamine-Lechani and Bahsoun, 2006; Teevan et al., 2005; Widyantoro et al., 1999) où une approche naïve consiste à considérer un profil de façon uni-dimensionnelle modélisant le centre d'intérêt général de l'utilisateur en intégrant des techniques de pondération des termes selon la formule du TF*IDF (Lieberman, 1997) ou selon l'approche probabiliste (Teevan et al., 2005). Par exemple, Teevan et al. (2005) définissent un score de personnalisation qui repose sur la construction d'un profil utilisateur à partir de ses jugements de pertinence. À l'inverse du principe probabiliste où l'information pertinente est incluse dans la collection de documents, Teevan et al. (2005) utilisent les répertoires personnels des utilisateurs, notés R , comme hypothèse de pertinence, comme illustré dans la Figure 2.7. À partir de l'hypothèse de pertinence développée par Teevan et al. (2005), les variables N et n_v présentées dans l'équation 2.4 et illustrées dans la partie gauche de la figure, sont respectivement transformées par le groupement des ensembles N et R comme suit : $N' = (N + R)$ et $n_{v'} = n_v + r_v$. Par conséquent, la formule du modèle probabiliste (Robertson et al., 1995) est revisitée en estimant le poids w_v d'un terme t_v de la façon suivante :

$$w_v = \log \frac{(r_v + 0.5)(N - n_v + 0.5)}{(n_v + 0.5)(R - r_v + 0.5)} \quad (2.16)$$

Le score de similarité $RSV(d_i, q_h)$ entre un document d_i et une requête q_h correspond au score personnalisé $perso(d_i, q_h, u_j)$ d'un document d_i pour un

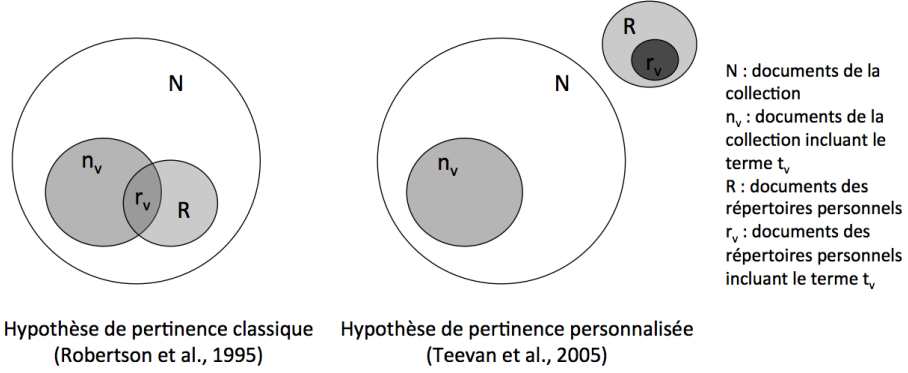


FIGURE 2.7: Comparaison de l'espace documentaire des hypothèses de pertinence traditionnelle (Robertson et al., 1995) et personnalisée (Morris and Horvitz, 2007)

utilisateur u_j est alors estimé ainsi :

$$RSV(d_i, q_h) = perso(d_i, q_h, u_j) = \prod_{t_v \in q_h} w_v \quad (2.17)$$

A l'inverse, d'autres approches considèrent une représentation multi-dimensionnelle où chaque dimension du profil représente un centre d'intérêt de l'utilisateur. Ainsi, les termes des documents visités et/ou des requêtes soumises sont agrégés à une des dimensions du profil de l'utilisateur (a) selon un seuil de similarité entre le document et le profil de l'utilisateur (Chen et al., 1998) ou (b) selon la polarité (positive ou négative) et la portée (court ou long terme) du jugement de pertinence (Widyantoro et al., 1999). Par exemple, Chen et al. (1998) modélisent le profil V par R vecteurs V_i correspondant respectivement aux R documents d_i jugés comme pertinents par l'utilisateur. Pour chaque nouveau document $d_{i'}$ sélectionné, la dimension V_i du profil la plus similaire à ce document $d_{i'}$ est mise à jour de la façon suivante :

$$V_i = V_i + V_{i'} \quad ; \quad V_i = \underset{V_i \in V}{\operatorname{argmax}} \operatorname{Sim}(V_i, V_{i'}) \quad (2.18)$$

avec $\operatorname{Sim}(V_i, V_{i'}) = \frac{V_i \cdot V_{i'}}{|V_i| \times |V_{i'}|}$

Seuls les m termes $t_v \in V_i$ dont les poids sont les plus élevés sont retenus pour mettre à jour la dimension V_i du profil V .

2. Les *approches connexionnistes* (Kim and Chan, 2003; Koutrika and Ioannidis, 2005; Micarelli and Sciarrone, 2004) reposent sur un réseau de termes. Par exemple, Micarelli and Sciarrone (2004) représentent un centre d'intérêt de l'utilisateur par un noeud annoté d'un terme, nommé planète, et relié à des termes satellites correspondant aux termes inclus dans les documents jugés pertinents où occure le terme planète. Dans ce même esprit, Koutrika and

Ioannidis (2005) proposent d'utiliser des opérateurs logiques entre les termes (Figure 2.8(a)), tandis que Kim and Chan (2003) hiérarchisent les termes, comme illustré dans la Figure 2.8(b), grâce à un algorithme de classification.

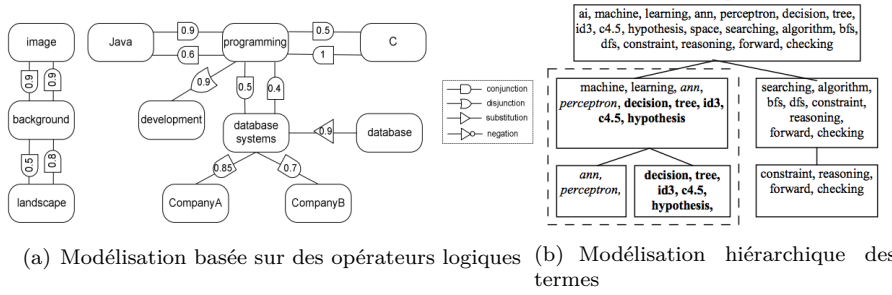


FIGURE 2.8: Modélisation des profils utilisateurs basée sur une approche connexionniste (Kim and Chan, 2003)

3. Les *approches conceptuelles* (Gauch et al., 2003; Leung et al., 2012; Liu et al., 2004; Sieg et al., 2007) reposent sur la représentation de concepts extraits à partir d'ontologies ou de hiérarchies de concepts, telles que ODP⁷ – *Open Directory Project* – qui est organisée en une hiérarchie de sujets associés à un ensemble de pages web caractéristiques. Le profil est alors modélisé sous la forme d'un réseau de concepts où les liens entre concepts respectent la topologie des liens des ressources conceptuelles.

Par exemple, Sieg et al. (2007) initialisent le profil utilisateur comme un ensemble de noeuds composés :

- (a) D'un score d'intérêt $IS(c)$ pour le concept c , initialisé à 1. Lors de la mise à jour du profil à partir des jugements de pertinence de l'utilisateur, chaque score d'intérêt $IS(c)$ est incrémenté par rapport à un score d'activation qui dépend de la similarité $Sim(d_i, c)$ du document sélectionné par l'utilisateur et du concept c . Le score d'activation $c.activation$ du concept c est estimé ainsi :

$$c.activation = IS(c) \cdot Sim(d_i, c) \quad (2.19)$$

Le score du concept c le plus élevé est propagé sur l'ensemble des concepts voisins de c selon le poids de la relation entre ces deux concepts. L'opération est effectuée de façon itérative.

- (b) De son vecteur moyen \vec{c} de termes associés inférés de l'agrégation des pages web $Docs(c)$ des sous-sujets $S(c)$ du concept c , comme présenté ci-dessous :

$$\vec{c} = \frac{\sum_{d \in Docs(c)} d}{|Docs(c)|} \quad (2.20)$$

avec $Sim(V_i, V_{i'}) = \frac{V_i \cdot V_{i'}}{|V_i| \times |V_{i'}|}$

7. <http://www.dmoz.org/>

2.2.1.2 Expertise de l'utilisateur

L'expertise d'un utilisateur est principalement reliée à l'ensemble de ses connaissances vis-à-vis d'un sujet particulier. Elle a été principalement étudiée dans le domaine du comportement utilisateur (Kang and Fu, 2010; White and Dumais, 2009). Les travaux de Xie and Joo (2012) ont mis en évidence le fait que l'expertise des utilisateurs infère sur le déroulement de la session (Kim and Allen, 2002) et par conséquent, doit être prise en compte comme un facteur contextuel du processus de RI. En effet, des précédents travaux ont montré que la technicité du vocabulaire utilisé (Allen, 1991; Hölscher and Strube, 2000; White and Dumais, 2009) et la longueur des requêtes (Kang and Fu, 2010; White and Dumais, 2009) peuvent être différentes entre un expert et un novice. De plus, un effet direct est perceptible au niveau des stratégies de recherche et de l'efficacité de la session (Hembrooke et al., 2005; White and Dumais, 2009), soulignant que les utilisateurs au faible niveau d'expertise sont limités dans la construction de lien sémantique autour du sujet de recherche et, par conséquent, n'approfondissent pas le sujet entièrement.

En RIS, l'expertise est également utilisée pour la détection d'acteurs d'experts d'un domaine. Nous identifions deux techniques principales :

1. *Les techniques basées sur les termes* consistent à estimer l'expertise à partir de la fréquence des termes des documents visités par l'utilisateur (Bharat and Mihaila, 2001; Demartini et al., 2009). Pour cela, les auteurs se basent sur une métrique statistique estimée à partir de la fréquence des termes, et plus particulièrement l'IDF analysant la distribution des termes dans la collection. Par exemple Bharat and Mihaila (2001) proposent d'estimer le score d'expertise $S_{Expert}(e_j)$ d'un utilisateur e_j selon une combinaison de trois scores :

$$S_{Expert}(e_j) = 2^{32} \cdot S_0 + 2^{16} \cdot S_1 + S_2 \quad (2.21)$$

où S_i correspond à un score basé sur les phrases des pages visitées qui contiennent $k - i$ termes de la requête, où k représente le nombre de termes dans la requête. Ce score dépend de deux facteurs : (a) $LevelScore(p)$ qui assigne un indicateur en fonction du type de la phrase (titre, en-tête, corps, ...) et $FullnessFactor(p, q_h)$ qui mesure le nombre de termes dans la phrase p qui recouvrent la requête q_h . Le score S_i est estimé ainsi :

$$S_i = \sum pLevelScore(p) \cdot FullnessFactor(p, q_h) \quad (2.22)$$

2. *Les techniques basées sur les modèles de langue* utilisent le vocabulaire de l'utilisateur pour construire son modèle de langue grâce à des distributions de probabilités multinomiales et identifier son niveau d'expertise par rapport à la requête (Balog et al., 2009; Pal and Counts, 2011; Petkova and Croft, 2006). Balog et al. (2009) identifient le niveau d'expertise d'un utilisateur grâce à la probabilité qu'un utilisateur candidat ca soit expert du sujet de la

requête q_h comme énoncé ci-dessous :

$$p(ca|q_h) = \frac{p(q_h|ca) \cdot p(ca)}{p(q)} \propto p(q_h|ca) \cdot p(ca) \quad (2.23)$$

Les auteurs représentent chaque utilisateur ca par une distribution multinomiale des termes issus du répertoire personnel des utilisateurs, tels que sa page web, ses emails ou ses articles. Un modèle de langue θ_{ca} est ensuite inféré permettant ainsi d'estimer la probabilité $p(q_h|ca)$ grâce à la formule du modèle de langue suivante :

$$\begin{aligned} p(q_h|\theta_{ca}) &= \prod_{t_v \in q_h} p(t_v|\theta_{ca})^{n(t_v, q_h)} \\ \text{avec } p(t_v|\theta_{ca}) &= (1 - \lambda)p(t_v|ca) + \lambda p(t_v) \end{aligned} \quad (2.24)$$

où $n(t_v, q_h)$ exprime le nombre d'occurrences du terme t_v dans la requête q_h , $p(t)$ est la probabilité du terme t_v dans le répertoire personnel de l'ensemble des utilisateurs. $p(t_v|ca)$ est la probabilité du terme t_v sachant le répertoire R de l'utilisateur. Cette probabilité est estimée à partir des documents d_i qui appartiennent au répertoire personnel permettant de relier les termes t_v au candidat ca grâce à la probabilité $p(d_i|ca)$ du document d_i sachant le candidat ca et la probabilité $p(t_v|d_i, ca)$ du terme t_v sachant le document d_i appartenant au répertoire du candidat ca :

$$p(t_v|ca) = \sum_{d_i \in R} p(t_v|d_i, ca) \cdot p(d_i, ca) \quad (2.25)$$

2.2.1.3 Rôle de l'utilisateur

De nombreuses taxonomies de rôles (Golder and Donath, 2004; Golovchinsky et al., 2009b; Welser et al., 2007) ont été proposées. Par exemple, Golder and Donath (2004) considèrent un réseau de forums, appelés Usenet, pour analyser les communautés d'information. A partir des rôles définis, ils mettent en évidence leurs stratégies ainsi que les liens entre les différents rôles. Egalement, Welser et al. (2007) analysent une situation de discussion afin de caractériser le rôle de "répondant" et de "discuteur" par une analyse statistique dans l'objectif de déterminer la signature du rôle. Ils utilisent ensuite une régression afin de confirmer le comportement du rôle. Dans un contexte de collaboration, les rôles (Golovchinsky et al., 2009b; Shah et al., 2010) sont définis afin d'assurer généralement la complémentarité et la diversité des compétences des utilisateurs (Sonnenwald et al., 2004), comme la paire d'expert de recherche ou de domaine.

Le Tableau 2.1 synthétise les rôles définis dans les trois taxonomies énoncées précédemment.

Selon une approche différente, de nombreuses contributions, principalement reliées au domaine de la RIS, se sont intéressées à l'apprentissage des rôles. Une

RIS		RIC
Golder and Donath (2004)	Welser et al. (2007)	Golovchinsky et al. (2009b)
Rôdeur “ <i>Lurker</i> ”	Répondant	Pairs
Provocateur “ <i>Troll</i> ”	Discuteur	Experts de domaine A/B
Débutant “ <i>Newby</i> ”		Expert/novice de recherche
Célébrité “ <i>Celebrity</i> ”		Expert/novice de domaine
Fanatique “ <i>Ranter</i> ”		Expert de recherche/domaine
		Prospecteur/Mineur

TABLE 2.1: Taxonomies des rôles en RIS et RIC

première catégorie de travaux est axée sur la structure des réseaux sociaux et des interactions entre les utilisateurs (Nowicki and Snijders, 2001; Kwak et al., 2010; Henderson et al., 2012). Par exemple, Nowicki and Snijders (2001) ont proposé une méthode d’apprentissage basée sur une analyse des graphes reposant sur des méthodes probabilistes génériques, à savoir l’approche de *Gibbs Sampling* (Geman and Geman, 1984). D’autre part, Henderson et al. (2012) se sont focalisés sur l’apprentissage de p rôles latents en se basant sur une méthode de transformation d’une matrice de caractéristiques des utilisateurs, où p est déterminé de façon non supervisée. L’idée principale est d’extraire le modèle de compression qui résume le comportement des utilisateurs dirigés par des rôles. Pour cela, les auteurs considèrent une matrice $V \in \mathcal{R}^{n \times f}$ où chaque élément $V(u_j, f_k)$ exprime la valeur de la caractéristique de comportement f_k pour l’utilisateur u_j . L’objectif est d’identifier la meilleure décomposition en valeurs singulières (SVD) qui satisfait la condition suivante :

$$\begin{aligned} & \underset{G, F}{\operatorname{argmin}} ||V - GF|| \\ & \text{sous la contrainte } G \geq 0 \text{ et } F \geq 0 \end{aligned} \quad (2.26)$$

où $G \in \mathcal{R}^{n \times p}$ correspond à la matrice qui connecte les utilisateurs aux p rôles latents et $F \in \mathcal{R}^{p \times f}$ relie les p rôles latents aux caractéristiques de comportement.

La deuxième catégorie de travaux s’intéresse à la détection des rôles par similarité ou différence des interactions entre les utilisateurs (McCallum et al., 2007). Ainsi, à partir d’un algorithme basé sur l’Allocation de Dirichlet Latente, appelé “*Author-Recipient-Topic* (ART)”, les auteurs proposent un modèle global permettant d’identifier les rôles utilisateur émergents dans un réseau de communication. Ce modèle est ensuite prolongé afin de mettre en évidence les rôles d’auteurs et de destinataires basés sur l’hypothèse que les rôles peuvent être fixes ou évoluer au cours du processus d’écriture des messages.

2.2.2 L’exploitation du profil utilisateur

Nous nous consacrons dans cette section à l’exploitation du profil utilisateur en distinguant les trois facteurs de centres d’intérêts, d’expertise et de rôle.

2.2.2.1 Intégration des centres d'intérêts de l'utilisateur

L'exploitation du profil, dans le cas de la prise en compte des centres d'intérêts, est orientée vers deux objectifs principaux : (a) l'accès à l'information textuelle (Agichtein et al., 2006; Larson et al., 2011; Liu et al., 2010; Radinsky et al., 2013; Sieg et al., 2007; Zhang et al., 2011) et (b) l'identification des principaux acteurs reliée aux intérêts des utilisateurs (Li et al., 2008; Newman and Girvan, 2004; Symeonidis et al., 2008).

En RIP, une des techniques utilisée repose principalement sur l'estimation de la pertinence des documents vis-à-vis de la requête et du profil de l'utilisateur (Sieg et al., 2007). Les auteurs proposent d'ordonner les documents à partir du concept c dont le score d'intérêt $IS(c)$ pour l'utilisateur u_j est le plus élevé de la façon suivante :

$$RSV(d_i|q_h, u_j) = IS(c) \cdot \text{sim}(q_h, d_i) \cdot \text{sim}(q_h, c) \cdot \gamma \quad (2.27)$$

où les scores de similarité sont estimés à partir des vecteurs de termes pondérés par la mesure TF*IDF et γ représente un coefficient de pondération.

D'autres travaux ne proposent pas d'étape intermédiaire de construction du profil de l'utilisateur mais injecte directement ses jugements de pertinence dans le modèle d'ordonnancement (Agichtein et al., 2006). Les auteurs proposent d'ordonner les documents en fonction du modèle de comportement des utilisateurs sur la base de leurs caractéristiques de comportement, telles que la longueur des requêtes ou le temps de visite par page. L'ordonnancement est effectué au travers d'un modèle d'apprentissage. Le poids $o(q_h, d_i, f_k, u_j)$ des caractéristiques f_k dans le modèle de l'utilisateur u_j est évalué dynamiquement par rapport au document d_i sélectionné et à la requête q_h associée en intégrant la tendance globale $C(f_k)$ des autres utilisateurs par rapport à cette même caractéristique f_k afin de réduire l'effet des actions bruitées de l'utilisateur :

$$o(q_h, d_i, f_k, u_j) = C(f_k) + \text{rel}(q_h, d_i, f_k, u_j) \quad (2.28)$$

où $\text{rel}(q_h, d_i, f_k, u_j)$ correspond à la valeur comportementale de l'utilisateur u_j vis-à-vis de la caractéristique f_k au dernier document d_i sélectionné pour la requête q_h .

Dans le même esprit, les approches collaboratives formulent l'hypothèse que la pertinence des documents repose sur les relations entre individus (Herlocker et al., 1999; Resnick et al., 1994). Du point du FCI, les documents sont ordonnés selon un score de prédiction $P(d_i, u_j)$ à partir de la préférence, traduite par une note $r_{u_j, i}$, des utilisateurs $u_j \in U$ qui partagent les mêmes intérêts, comme détaillé ci-dessous (Resnick et al., 1994) :

$$P(d_i, u_j) = \frac{\sum_{u_{j'} \in U} (r_{u_{j'}, i} - \bar{r}_j) \cdot w_{j', j}}{\sum_{u_{j'} \in U} |w_{j', j}|} \quad (2.29)$$

$$\text{avec } w_{j', j} = \frac{\sum_{d_i \in \mathcal{D}} (r_{u_j, i} - \bar{r}_j)(r_{u_{j'}, i} - \bar{r}_{j'})}{\sqrt{\sum_{d_i \in \mathcal{D}} (r_{u_j, i} - \bar{r}_j)^2} \sqrt{\sum_{d_i \in \mathcal{D}} (r_{u_{j'}, i} - \bar{r}_{j'})^2}}$$

où $w_{j,j'}$ correspond à la corrélation entre deux utilisateurs u_j et $u_{j'}$ et $\bar{r}_{j'}$ exprime la note moyenne des documents jugés par l'utilisateur $u_{j'}$.

Une extension de ces approches d'ordonnancement réside dans l'intégration de méta-données sociales, telles que les annotations, des jugements de pertinence des utilisateurs (Zhang et al., 2011) ou de la confiance accordée à la relation entre les utilisateurs (Ma et al., 2009).

Une autre exploitation intéressante consiste en l'identification d'utilisateurs ou de communautés partageant les mêmes centres d'intérêts. Pour cela, Li et al. (2008) utilisent également les annotations sociales comme source d'évidence et facteur de partage de centres d'intérêts entre utilisateurs tandis que Newman and Girvan (2004) considèrent la structure du réseau. Plus particulièrement, les auteurs formulent l'hypothèse qu'un utilisateur central, détecté à partir de la mesure de *betweenness* (Freeman, 1977), est un utilisateur à l'interconnection de plusieurs communautés et que supprimer ces acteurs du réseau permet de faire émerger les communautés d'intérêts du réseau.

2.2.2.2 Intégration de l'expertise de l'utilisateur

La plupart des travaux du domaine s'appuient sur le score d'expertise des utilisateurs comme facteur de pertinence afin d'ordonner les documents (Bharat and Mihaila, 2001; Ben Jabeur et al., 2010; Kirsch et al., 2006). L'objectif est de considérer un ensemble d'utilisateurs afin d'identifier les documents jugés, visités ou écrits par les experts du domaine, appelés également autorités. Par exemple, Bharat and Mihaila (2001) proposent d'identifier les sujets populaires à partir des pages visitées et des requêtes soumises par un ensemble d'experts. Pour cela, les auteurs construisent un graphe bi-parti où les relations entre les experts e_j et les documents d_i sont pondérées par un score $S_{Edge}(e_j, d_i)$ de la façon suivante :

$$S_{Edge}(e_j, d_i) = \begin{cases} S_{Expert}(e_j) \cdot \sum_{t_v \in q_h} occ(t_v, d_i) & \text{si } \exists t_v \in q_h, occ(t_v, d_i) > 0 \\ 0 & \text{sinon} \end{cases} \quad (2.30)$$

où $S_{Expert}(e_j)$ correspond au score d'expertise de l'expert e_j estimé à partir de la correspondance entre les termes des requêtes soumises et ceux des URLs associées. $occ(t_v, d_i)$ exprime le nombre de phrases dans le document d_i qui contiennent le terme t_v . Le score $S_{doc}(d_i)$ d'un document d_i est ensuite estimé par la somme des poids des ses relations, comme indiqué ci-dessous :

$$S_{doc}(d_i) = \sum_{e_j \in O(d_i)} S_{Edge}(e_j, d_i) \quad (2.31)$$

où $O(d_i)$ correspond à la fonction qui retourne les experts reliés au document d_i dans le graphe.

Dans certains travaux orientés vers la recherche bibliographique (Ben Jabeur et al., 2010; Kirsch et al., 2006), les auteurs proposent un modèle d’ordonnement social des documents en se basant sur un réseau bibliographique pondéré composé d’auteurs et d’articles scientifiques. Par exemple, Ben Jabeur et al. (2010) utilisent les interactions sociales, telles que la citation, l’annotation, la relation co-auteur, afin de privilégier les documents écrits, visités ou annotés par des autorités/experts du domaine. Pour cela, les auteurs combinent un score de similarité textuelle $RSV(q_h, d_i)$ avec un score d’importance sociale $Imp(d_i)$ du document qui agrège le score de centralité des auteurs reliés au document d_i :

$$RSV(q_h, d_i) = \alpha RSV(q_h, d_i) + (1 - \alpha) Imp(d_i) \quad (2.32)$$

2.2.2.3 Intégration du rôle de l’utilisateur

La plupart des travaux liés à l’exploitation du rôle dans un modèle d’ordonnement se sont consacrés à l’appariement des utilisateurs vis-à-vis de ce rôle dans un contexte social (Demartini, 2007; Ben Jabeur et al., 2012; Kwak et al., 2010). Par exemple, Zhang et al. (2007) proposent une extension de l’algorithme PageRank qui repose sur l’hypothèse que si un utilisateur est central dans un réseau, il est expert. Ainsi, l’algorithme, appelé *ExpertiseRank*, propage les poids $ER(u_j)$ des noeuds utilisateurs u_j dans le réseau jusqu’à convergence. Dans le contexte particulier d’un système de questions-réponses (RIS), les auteurs considèrent que l’utilisateur u_j a répondu à n utilisateurs $u_{j'} \in U$. Le score de l’utilisateur u_j est alors estimé de la façon suivante :

$$ER(u_j) = (1 - d) + d \sum_{u_{j'} \in U} \frac{ER(u_{j'})}{C(u_{j'})} \quad (2.33)$$

avec $C(u_{j'})$ représente le nombre total d’utilisateurs qui aident $u_{j'}$. d est le coefficient de pondération.

De plus, certains auteurs (Pickens et al., 2008; Shah et al., 2010) du domaine de la RIC se sont basés sur des rôles prédéfinis dans la taxonomie de (Golovchinsky et al., 2009b) afin de modéliser des modèles d’ordonnement collaboratifs des documents. Brièvement, ces modèles s’appuient sur les objectifs des rôles, par exemple “favoriser la diversité”, et exploitent les techniques de RI, dans notre exemple la reformulation de requêtes, afin de guider le collaborateur dans sa tâche. Ces modèles font l’objet d’une explication plus détaillée dans le chapitre suivant (Section 3.5.2.3), car ils constituent le fondement de nos travaux.

2.2.3 Synthèse des approches de la RI orientée utilisateurs

Le Tableau 2.2 recense l’ensemble des travaux de la RI orientée utilisateurs qui ont modélisé ou exploité les dimensions de centres d’intérêts, d’expertise et de rôle, intrinsèques aux utilisateurs.

Modélisation du profil	Centres d'intérêts	Approches basées sur les termes	(Chen et al., 1998; Lieberman, 1997; Teevan et al., 2005; Low et al., 2011; Brosseau-Villeneuve et al., 2014)
		Approches connexionnistes	(Kim and Chan, 2003; Koutrika and Ioannidis, 2005; Micarelli and Sciarrone, 2004)
		Approches conceptuelles	(Gauch et al., 2003; Leung et al., 2012; Liu et al., 2004; Sieg et al., 2007)
	Expertise	Approches basées sur les termes	(Bharat and Mihaila, 2001; Demartini, 2007; Mimno and McCallum, 2007)
		Approches basées sur le modèle de langue	(Balog et al., 2009; Pal and Counts, 2011)
	Rôle	Taxonomies	(Golder and Donath, 2004; Golovchinsky et al., 2009b; Welser et al., 2007)
Exploitation du profil		Apprentissage des rôles	(Henderson et al., 2012; Kwak et al., 2010; Nowicki and Snijders, 2001)
	Centres d'intérêts	Ordonnancement des documents	(Agichtein et al., 2006; Liu et al., 2010; Resnick et al., 1994; Sieg et al., 2007)
		Identification d'utilisateurs	(Li et al., 2008; Newman and Girvan, 2004)
	Expertise	Ordonnancement des documents	(Bharat and Mihaila, 2001; Ben Jabeur et al., 2010; Kirsch et al., 2006)
	Rôle	Ordonnancement des documents	(Ben Jabeur et al., 2012; Pickens et al., 2008; Shah et al., 2010; Zhang et al., 2007)

TABLE 2.2: Synthèse des travaux de la RI orientée utilisateur

Résumé

Ce chapitre présente, dans un premier temps, les notions relatives à la RI ainsi que les principaux travaux de la littérature. Dans un second temps, nous avons axé notre état de l'art sur la composante essentielle d'un processus de RI adapté à la dimension utilisateur. Nous nous sommes donc intéressés à la prise en compte du facteur individu dans le FIC, la RIS et RIC selon trois dimensions : les centres d'intérêts, l'expertise et le rôle des utilisateurs. Les deux enjeux liés à la modélisation et l'exploitation du profil de l'utilisateur ont été abordés.

Dans le chapitre suivant, nous présentons la RIC qui, en complément des approches précédentes, considère un environnement multi-utilisateurs. Nous nous focalisons particulièrement sur les principales notions liées à la collaboration, aux différentes approches et aux cadres d'évaluation de ce domaine.

Chapitre 3

Recherche d'information collaborative

Introduction

Des études cognitives (Foster, 2006; Talja, 2002) ont souligné le besoin de collaboration dans de nombreux domaines d'applications, tels que le domaine médical (Morris and Morris, 2011; Reddy and Spence, 2002; Reddy et al., 2002), académique (Foster, 2006; Large et al., 2002) ou du e-Discovery (Attfield et al., 2010; Conrad, 2007). De plus, l'émergence du web 2.0 a donné une nouvelle dimension à la RI en positionnant l'utilisateur au centre du processus de recherche (Amer-Yahia et al., 2007). Les utilisateurs ont alors la possibilité d'interagir ou d'organiser leurs actions, par le biais de réseaux sociaux ou d'outils de collaboration en ligne.

Plus proche de nos travaux, la collaboration se manifeste également au sein d'un processus de RI (Morris, 2013), notamment dans le cas de tâches complexes ou exploratoires (Singer et al., 2012) où la considération d'un environnement de recherche individuel peut atteindre ses limites (Denning and Yaholkovsky, 2008; London, 1995). La collaboration repose sur la complémentarité des compétences des individus (Sonnenwald et al., 2004; Spence et al., 2005) et la coordination entre les utilisateurs. Elle permet alors d'améliorer la qualité des résultats de recherche par l'effet synergique de la collaboration (Shah and González-Ibáñez, 2011b) et l'utilisation de l'intelligence collective du groupe (Evans and Chi, 2010; Twidale et al., 1997). Dans ce contexte, la RIC est définie selon Hansen and Järvelin (2005) de la façon suivante :

"CIR is an information access activity related to a specific problem solving activity that, implicitly or explicitly, involves human beings interacting with other human(s) directly and/or through texts (e.g., documents, notes, fi-

gures) as information sources in an work task related information seeking and retrieval process either in a specific workplace setting or in a more open community or environment."

Des travaux (Foley, 2008; Morris and Horvitz, 2007) ont été présentés dans le domaine de la RIC dont l’objectif est de résoudre un besoin en information partagé par un ensemble d’utilisateurs. Joho et al. (2009) proposent trois axes principaux de support à la collaboration :

- *Les interfaces* (Morris and Horvitz, 2007; Shah and González-Ibáñez, 2011b) offrent un espace de travail individuel et partagé permettant d’interagir entre collaborateurs et d’avoir une visibilité sur les actions de la communauté.
- *Les techniques* (Foley, 2008) de "*relevance feedback*", de classification, de modélisation de profil, de fusion de résultats sont revisitées pour intégrer une composante collaborative dans les modèles.
- *Les modèles* (Pickens et al., 2008; Shah et al., 2010) considèrent un besoin en information collectif et intègrent, dans les algorithmes d’appariement, différentes sources d’évidence telles que la structure du réseau de collaboration, les rôles des collaborateurs, les jugements de pertinence et les historiques de recherche.

Dans ce chapitre, nous relierons dans la section 3.1 le concept de la collaboration au domaine de la RI. La section 3.2 présente les principaux concepts de la RIC. Nous détaillons les modèles génériques de la RIC dans la section 3.3 et illustrons les principaux domaines d’application dans la section 3.4. Nous présentons, dans la section 3.5, les différentes approches de la RIC proposées dans la littérature. Enfin, nous décrivons dans la section 3.6 les contraintes imposées par la collaboration sur la définition des cadres d’évaluation pour la RIC.

3.1 Recherche d’information et collaboration

La RIC repose sur la notion de collaboration (Gray, 1989) entre les utilisateurs qui est motivée par un besoin en information partagé et l’objectif de produire un résultat qui dépasse l’agrégation des résultats individuels (Shah and González-Ibáñez, 2011b). Afin d’optimiser le processus de collaboration, le contexte multi-utilisateurs implique la mise en place de paradigmes qui permettent de structurer les actions des collaborateurs.

Cette section est un préambule à la présentation des travaux du domaine de la RI afin de définir la notion de collaboration (Section 3.1.1), d’énoncer les motivations (Section 3.1.2) et d’introduire les paradigmes (Section 3.1.3).

3.1.1 Notion de collaboration

Définition. Issue du latin “cum” et “laborare”, la notion de collaboration peut être synthétisée par l’action de “travailler ensemble”. De nombreux travaux

(Gray, 1989; London, 1995; Wood and Gray, 1991) se sont intéressés à définir cette notion dont la plus robuste, selon London (1995), est la suivante :

“A process through which parties who see different aspects of a problem can constructively explore their differences and search for solutions that go beyond their own limited vision of what is possible.” (Gray, 1989)

Plus récemment, une autre définition proposée par Thomson and Perry (2006) résume l'ensemble des définitions ainsi :

“Collaboration is a process in which autonomous actors interact through formal and informal negotiation, jointly creating rules and structures governing their relationships and ways to act or decide on the issues that brought them together; it is a process involving shared norms and mutually beneficial interactions.” (Thomson and Perry, 2006)

La complétude de cette définition permet d'identifier cinq points clés de la collaboration (Thomson and Perry, 2006), qui sont également abordés dans d'autres travaux (Assent, 2011; Surowiecki, 2004) :

1. La *gouvernance* : la prise de décision à propos des règles de coordination de la collaboration doit être conjointe à l'ensemble des membres du groupe.
2. L'*administration* : le système de collaboration doit permettre la transparence des rôles de chacun et proposer un moyen de communication.
3. L'*autonomie* : la collaboration doit permettre un juste équilibre entre l'intérêt collectif de satisfaire le besoin partagé et l'intérêt individuel du participant à satisfaire son propre besoin informationnel.
4. La *mutualité* : la collaboration repose sur les différences des utilisateurs pour arriver à bénéficier de chacun.
5. Les *normes* : la collaboration repose sur la confiance et des normes de réciprocité.

Selon une vision terminologique, la définition de la collaboration est proche de celle des notions de coopération et de coordination. Shah (2012) décrit ainsi cinq niveaux encapsulés qui permettent d'avoir une vision plus fine de la collaboration. Ces niveaux sont illustrés dans la Figure 3.1 où chaque niveau, du plus granulaire au plus large est présenté ci-dessous : (a) *le niveau de communication* qui permet aux utilisateurs d'échanger des informations afin d'établir une relation productive ; (b) *le niveau de contribution* qui permet à un individu d'aider les autres afin de satisfaire leurs objectifs individuels ; (c) *le niveau de coordination* qui favorise le partage des ressources, des responsabilités et des objectifs ; (d) *le niveau de coopération* qui permet une relation basée sur des intérêts similaires, tout en considérant les tâches de négociation des rôles ou du partage des ressources ; (e) *le niveau de collaboration* qui permet aux utilisateurs d'interagir dans un objectif commun en considérant les différentes compétences et les divergences de chacun au niveau de la vision du domaine. Les utilisateurs sont complémentaires et cherchent des solutions communes.

Dans ce contexte, la collaboration peut être perçue comme une notion multidimensionnelle (Capra et al., 2010; Golovchinsky et al., 2009a,b), caractérisée par les aspects suivants :

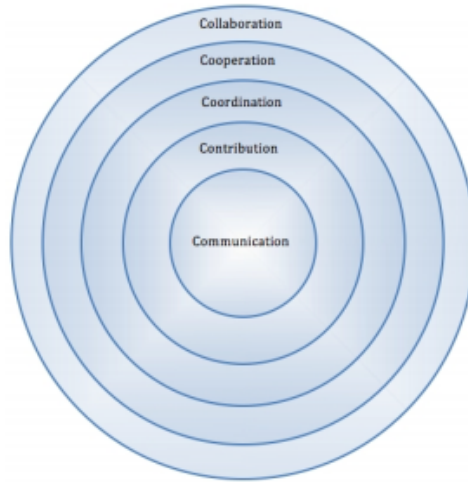


FIGURE 3.1: Les cinq niveaux de la collaboration selon Shah (2012)

1. **L'intention.** Elle peut être *explicite* ou *implicite*, suivant si les utilisateurs formulent une réelle intention de collaborer ou non. La collaboration *implicite* fait plutôt appel à la problématique du FCI ou de la RIS où les documents retournés à un utilisateur sont influencés par les jugements de pertinence d'autres utilisateurs, sans en avoir explicitement conscience. A l'inverse, la collaboration *explicite* est caractérisée par une situation où les utilisateurs expriment le besoin de s'unir pour satisfaire un objectif commun.
2. **Le niveau de médiation.** La collaboration peut être supportée à deux niveaux différents. D'une part, la *médiation utilisateur* s'appuie particulièrement sur des interfaces adaptées à la collaboration, servant uniquement d'outil de communication, d'interaction et de partage d'information entre les utilisateurs (Hartmann et al., 2009; Smeaton et al., 2006). D'autre part, la *médiation système* s'intéresse à l'accompagnement algorithmique de la collaboration en proposant des modèles d'ordonnancement qui gèrent le contexte multi-utilisateurs (Foley and Smeaton, 2009; Pickens et al., 2008; Shah et al., 2010).
3. **La concurrence.** Elle peut être *synchrone* ou *asynchrone*, suivant si les utilisateurs cherchent simultanément ou non. La plupart des modèles d'ordonnancement présentés à ce jour dans la littérature considèrent une collaboration *synchrone* (Foley and Smeaton, 2009; Pickens et al., 2008; Shah et al., 2010).
4. **La localisation.** Les collaborateurs peuvent collaborer de différentes sortes : a) dans la même pièce, sur un même ordinateur ou sur des ordinateurs séparés ; c'est une collaboration *proche –collocated–* ou b) dans des pièces différentes caractérisant une collaboration *éloignée –remote–*.
5. **L'implication.** Une collaboration est *active*, respectivement *passive*, si les

utilisateurs annotent, commentent et/ou évaluent les documents, respectivement si les utilisateurs naviguent simplement dans le web (Golovchinsky et al., 2009a,b).

3.1.2 Motivations

Fidel et al. (2004) ont démontré que la mise en place de la collaboration dans un processus de RI ajoute une charge cognitive supplémentaire par rapport à un contexte de recherche individuelle. Cependant, pour de nombreuses tâches, la collaboration est nécessaire et bénéfique (Denning and Yaholkovsky, 2008). Nous présentons ci-dessous les motivations qui poussent les individus à collaborer, malgré cette charge cognitive supplémentaire.

Tâche de collaboration imposée. Une des motivations les plus naïves de la RIC repose sur le fait que les utilisateurs collaborent car une tâche de travail en groupe leur a été imposée. C'est souvent le cas dans le domaine universitaire (Hyldegård, 2006), où les étudiants collaborent pour réaliser un projet ou une synthèse bibliographique qui impliquent une RI active sur le web. Dans ces cas, la collaboration n'est pas initiée par les utilisateurs eux-mêmes qui peuvent être caractérisés par une faible motivation. Il est donc nécessaire de créer des outils ou modèles adaptés et permettant l'individualisation dans le travail pour leur permettre de retrouver une certaine autonomie et indépendance, malgré la contrainte de la collaboration imposée.

Besoin en information commun. Lors d'une tâche non imposée, le fondement de la collaboration réside dans le besoin en information que partagent les collaborateurs. Teevan et al. (2009) ont étudié les caractéristiques des utilisateurs au sein d'un même groupe et ont identifié, que dans le cas de tâches complexes nécessitant réellement de la collaboration (London, 1995), les utilisateurs ont des profils traduisant des centres d'intérêts similaires. Un peu en marge de la RIC définie par Hansen and Järvelin (2005), les systèmes de questions-réponses donnent également un fondement à la collaboration, bien qu'elle soit implicite. Dans ces systèmes, le besoin en information émane d'un utilisateur qui requiert l'expertise d'un autre utilisateur, caractérisé par des mêmes intérêts, pour répondre à son propre besoin initial (Morris and Teevan, 2009). Les réponses de l'expert peuvent être constituées de références bibliographiques qui résultent d'un processus de RI.

Connaissances individuelles insuffisantes. Dans une étude, Morris and Teevan (2009) ont mis en évidence que certains utilisateurs font appel à des systèmes de questions-réponses lorsqu'ils n'ont pas trouvé de solution à leur besoin en information en cherchant de façon individuelle. Bien qu'étant une tâche particulière en marge de la RIC, ce constat permet de souligner le fait que pour certaines

tâches, notamment lors de tâches complexes ou exploratoires, les utilisateurs ont tout intérêt à collaborer lors du processus de RI afin d'obtenir de meilleurs résultats. En effet, le manque de connaissances d'un unique utilisateur pour ce type de besoin en information génère un besoin de collaboration qui permet de connecter des utilisateurs aux compétences et connaissances complémentaires (Sonnenwald et al., 2004; Spence et al., 2005). L'avantage de la collaboration est qu'elle permet de couvrir l'ensemble des aspects du besoin tout en satisfaisant le bénéfice mutuel des individus. Ainsi, les utilisateurs peuvent accroître leur connaissance du sujet et trouver un rapport apprentissage/contribution qui motive l'utilisateur à collaborer. C'est ce que Morris (2013) et Gardner (2005) ont appelé le bénéfice mutuel ou également, une situation gagnante-gagnante – *Win-Win situation*–. De plus, la collaboration bénéficie d'un effet synergique, caractérisé par le fait que le résultat du groupe dépasse l'agrégation des résultats individuels (Shah and González-Ibáñez, 2011b).

Dimension sociale de la collaboration. La dimension sociale de la collaboration (Morris and Teevan, 2009) est également non négligeable. Certaines études (Evans and Chi, 2010; Morris and Teevan, 2009) ont montré que la partie affective, liée à la reconnaissance sociale d'un individu, impacte sur la qualité des résultats lors d'une session de recherche. La collaboration permet à un utilisateur d'étendre ou de maintenir son réseau social, mais également de donner de la visibilité à ses connaissances et/ou compétences (Ackerman and Palen, 1996; Beenen et al., 2004).

3.1.3 Paradigmes de la RIC

Dans cette partie, nous détaillons les trois paradigmes sur lesquels repose la collaboration dans un processus de RI (Foley and Smeaton, 2010; Morris et al., 2010). Bien que chacun d'eux soit basé sur des hypothèses différentes, ces trois paradigmes permettent d'éviter la redondance entre les actions des utilisateurs et d'optimiser la collaboration aussi bien en termes d'efficacité que de diversité des résultats.

3.1.3.1 La division du travail

La division du travail –*Division of labour*– est le paradigme le plus connu. Il a pour objectif de séparer le travail entre les utilisateurs afin d'éviter la redondance d'exécution de leurs actions (Foley, 2008; Poltrock et al., 2003) et de gagner du temps dans la réalisation de la tâche de recherche (Zeballos, 1999). Pour répondre à ce problème lors d'une session de RI, les interfaces ou les techniques de RI proposées peuvent s'appuyer sur quatre types de division du travail (Kelly and Payne, 2013) :

1. *Les approches basées sur la communication* constituent la forme la plus simple de collaboration où les utilisateurs échangent des informations sur leur recherche et le contenu des informations visitées afin de coordonner leur travail. La communication peut se faire de manière naturelle, guidée par la stratégie “*diviser pour régner*” –*divide and conquer*– (Morris et al., 2008), ou par le biais de messageries instantanées (Shah and González-Ibáñez, 2011a).
2. *Les approches basées sur les interfaces utilisateur* consistent à diviser les ensembles documentaires (Foley and Smeaton, 2009; Morris and Horvitz, 2007) au moyen d’interfaces afin de garantir aux collaborateurs d’examiner des listes de documents différentes. En effet, les collaborateurs explorant le même besoin en information proposent des requêtes reformulées qui peuvent être similaires générant des résultats fortement similaires également.
3. *Les approches basées sur les algorithmes* utilisent les données générées par les collaborateurs afin de proposer des ordonnancements adaptés à leurs jugements de pertinence (Foley and Smeaton, 2009). Les coûts de coordination pour les utilisateurs sont très faibles, leur permettant de se consacrer uniquement à la recherche d’informations pertinentes. Deux limites majeures peuvent être identifiées : 1) en favorisant la prise en compte des actions des utilisateurs, la distribution des résultats n’est pas toujours en adéquation avec les connaissances des utilisateurs et 2) ces algorithmes ne redistribuent pas les documents sauvegardés, alors que les utilisateurs peuvent être amenés à les revoir.
4. *Les approches basées sur les rôles* font appel à une division des tâches (Foley and Smeaton, 2010; Pickens et al., 2008; Shah et al., 2010) de collaboration à travers l’attribution de rôles où les utilisateurs agissent selon des objectifs distincts. Par exemple, Pickens et al. (2008) ont considéré des scénarios de recherche où un utilisateur s’intéresse à la diversité des documents alors que son collaborateur se focalise sur les documents les plus pertinents.

3.1.3.2 Le partage des connaissances

Le partage des connaissances –*Sharing of knowledge* est une dimension fondatrice de la collaboration (Yao et al., 1999), surtout dans le cas où la collaboration repose sur une tâche de RI (Hansen and Järvelin, 2005). Le partage des connaissances fait référence à l’action de transmettre les informations entre collaborateurs pendant une activité de groupe (Foley and Smeaton, 2009). Il permet à un utilisateur de bénéficier des compétences complémentaires des membres du groupe (Foley and Smeaton, 2010) et peut être supporté de différentes façons :

- Dans un contexte de *médiation utilisateur*, le partage des connaissances peut être réalisé par le biais de la communication (Lay et al., 1988) ou d’espace de travail partagé (Lay et al., 1988; Poltrock et al., 2003).
- D’un point de vue *médiation système*, ce sont plutôt les jugements de pertinence des utilisateurs qui vont permettre d’influencer sur l’ordonnancement de tous les membres du groupe (Foley and Smeaton, 2009), dans le sens où

les jugements de pertinence permettent de construire l'intelligence collective (Harman, 1992) et de satisfaire le besoin en information commun.

3.1.3.3 La conscience de l'environnement collaboratif

La conscience de l'environnement collaboratif –*Awareness*–, c'est-à-dire des collaborateurs et de leurs actions, est une partie intégrante du processus de collaboration afin de permettre aux individus de se coordonner (Dourish and Bellotti, 1992; Gaver, 1991). Ces derniers définissent la notion d'*awareness* ainsi :

“Awareness is an understanding of the activities of others, which provides a context for your own activity. This context is used to ensure that individual contributions are relevant to the group’s activity as a whole, and to evaluate individual actions with respect to group goals and progress. The information, then, allows groups to manage the process of collaborative working.” (Gaver, 1991)

Ce concept est prédominant dans la définition des caractéristiques des interfaces homme-machine (Rodriguez Perez et al., 2011; Morris and Horvitz, 2007). Une taxonomie a été proposée par Liechti and Sumi (2002), qui distinguent quatre types de conscience :

- La *conscience du groupe* consiste à fournir aux utilisateurs des informations à propos des statuts et des activités des membres du groupe.
- La *conscience de l'espace de travail* a pour objectif de fournir un espace de travail partagé permettant aux collaborateurs de discuter des résultats de leur recherche et de produire un résultat commun.
- La *conscience du contexte*, plutôt reliée au domaine d'application, consiste à déterminer ce qui est utile au groupe et les objectifs de la tâche.
- La *conscience périphérique* réfère à l'ensemble des informations en marge du processus, qui diffère de ce que les utilisateurs sont en train de regarder ou faire.

3.2 Concepts de base

3.2.1 Collaborateurs

L'élément fondateur de la collaboration est le groupe d'individus qui interagissent pour satisfaire un besoin en information commun. Dans un premier temps, nous nous intéressons au groupe dans sa globalité afin d'en étudier les principales caractéristiques. Ensuite, nous présentons quelques propriétés sur l'élément atomique de ce groupe de collaboration, qu'est l'individu.

3.2.1.1 Le groupe de collaboration

Selon (Morris and Teevan, 2009) et Teevan et al. (2009), les groupes de collaborateurs peuvent être analysés selon deux axes :

1. La *longévité* analyse le facteur temps. Un groupe basé sur des relations à court-terme est considéré comme un groupe de tâches –*task-based group*– où les individus partagent un objectif commun avec des recherches sur une courte période, telles que la planification d’un voyage ou des recherches transactionnelles. A contrario, un groupe avec des relations à long-terme est perçu comme un groupe de caractéristiques –*trait-based group*– où les individus partagent des intérêts à long-terme, soumettent généralement le même type de requêtes ou possèdent des informations pertinentes pour les autres personnes du groupe.
2. Le *mode d’identification* concerne la façon dont le groupe est identifié. Si le groupe partage des informations pour une même tâche, le mode est dit *explicite*. Dans ce contexte, les groupes utilisent des outils de collaboration adaptés à leur scénario de recherche permettant des interactions explicites entre les utilisateurs. A contrario, il est dit *implicite* lorsque la tâche commune est inférée du fait que les utilisateurs sont en train de réaliser les mêmes actions ou semblent se diriger vers le même résultat à partir d’un besoin similaire. Ces groupes de collaboration peuvent être déduits, par exemple, des données géographiques, démographiques, des historiques de recherche et de l’utilisation de certains sites internet à sujets spécialisés.

3.2.1.2 Le collaborateur

Les collaborateurs sont les éléments centraux d’un processus de collaboration et ont été analysés selon différents aspects. Une première catégorie de travaux s’est intéressée particulièrement au comportement de l’utilisateur en situation de recherche (Evans and Chi, 2010; Hyldegård, 2006), incluant la nature de son besoin de collaboration (Morris and Teevan, 2009) et sa stratégie de recherche (Capra, 2013). Parmi ces aspects, un élément est particulièrement intéressant dans le processus de collaboration car il permet aussi bien de régir les stratégies de collaboration que de garantir le paradigme de la division du travail, détaillé dans la section 3.1.3.1. Il s’agit du rôle des utilisateurs, présenté dans la taxonomie de Golovchinsky et al. (2009b) :

1. Le rôle de *pairs* où les collaborateurs agissent indépendamment et combinent leurs résultats manuellement. Les rôles sont donc symétriques, permettant aux collaborateurs de coordonner librement leurs actions. Foley and Smeaton (2009) ont proposé une médiation système pour ce type de rôles où les actions des utilisateurs sont exploitées dans le modèle par le biais des jugements de pertinence.
2. La paire de rôles *Expert d’un domaine A / Expert d’un domaine B* où les collaborateurs ont des domaines d’expertise différents. Les ordonnancements

peuvent donc être adaptés à ces domaines différents afin de garantir une meilleure couverture du besoin.

3. La paire de rôles *Expert de recherche / Novice en recherche* ou *Expert de domaine / Novice du domaine*, référençant le degré d’expertise des collaborateurs par rapport à un domaine ou aux outils de recherche. Pour des niveaux différents en termes d’expertise de recherche, un utilisateur peut être amené à utiliser la fonctionnalité de recherche avancée tandis que son collaborateur utilise plutôt l’interface de base. Dans le cas de niveaux d’expertise de domaine différents, la médiation se base plutôt au niveau algorithmique, comme dans la paire de rôles précédente.
4. La paire de rôles *Expert de recherche / Expert du domaine* où la contribution des deux utilisateurs au processus de recherche est asymétrique. L’expert de recherche sait comment sélectionner les documents et formuler les requêtes, mais peut juger les documents de façon rudimentaire sur la base de la description des besoins de l’expert du domaine. Ce dernier va être plus performant dans les jugements de pertinence des documents retrouvés par l’expert en recherche. Il peut les évaluer de façon précise.
5. La paire de rôles *Prospecteur / Mineur* où le *Prospecteur* a pour objectif de favoriser la diversité dans les espaces de recherche alors que le *Mineur* s’intéresse plus en profondeur à la pertinence des documents et explore ceux qui sont au coeur du besoin en information. Une médiation algorithmique a été proposée pour cette paire de rôles (Pickens et al., 2008). Nous la détaillons en section 3.5.2.3.

Indépendamment de la taxonomie, Shah et al. (2010) ont proposé une nouvelle paire de rôles, reposant sur les rôles de *Rassembleur –Gatherer–* et de *Surveilleur –Surveyor–*. La tâche du premier est d’identifier rapidement les documents pertinents tandis que celle du second est de sélectionner des documents en garantissant une diversité thématique afin d’avoir une connaissance plus large du sujet. Celle-ci lui permet ainsi de comprendre et d’ajuster les stratégies de recherche durant la session.

3.2.2 La tâche collaborative

La tâche de collaboration correspond à l’objectif fixé par les utilisateurs qui doit être résolu pendant le processus de recherche. Morris (2013) a réalisé un sondage afin d’identifier les habitudes de collaboration et plus particulièrement les tâches réalisées. Le Tableau 3.1 présente les tâches les plus communes, soulignant une prédominance pour celles de planification de voyage et d’achats en ligne où la collaboration est souvent utilisée pour demander des conseils par rapport à un produit (besoin transactionnel), pour atteindre un site de commerce en particulier (besoin navigationnel) ou encore l’exploration d’informations relatives à un objectif (besoin informationnel).

Nous pouvons ainsi dégager deux catégories de tâches de recherche particulièrement traitées dans un contexte collaboratif (Capra et al., 2011; Morris, 2013) :

Tâche de collaboration	Fréquence
Planification de voyage	27.5%
Achats en ligne	25.7%
Recherche bibliographique	20.2%
Recherche d'information technique	16.5%
Identifications de faits	16.5%
Planification d'évènements sociaux	12.8%
Recherche d'information médicale	6.4%
Immobilier	6.4%

TABLE 3.1: Statistiques sur la fréquence de collaboration (Morris, 2013)

- Les tâches “*fact-finding*” relèvent d’un besoin spécifique sur une information particulière (Jansen et al., 2008; Kellar et al., 2007; Malik et al., 2007). On dénombre les tâches relatives à l’identification de faits, comme la détection d’évènements ou de lieux, ou d’une information technique, comme la notice d’un appareil électroménager. Dans ces cas là, la collaboration peut être utilisée en termes de division du travail pour un gain de temps, par exemple.
- Les tâches exploratoires (Singer et al., 2012) relèvent d’un besoin plus large et plus complexe où plusieurs aspects et interprétations peuvent être traités. Des exemples de tâches exploratoires abordées en RIC sont la planification d’un voyage ou encore la résolution d’un besoin médical ou bibliographique.

3.2.3 Le processus de RIC

Le processus de RIC est défini comme un ensemble d’actions successives où les collaborateurs interagissent entre eux mais également avec le système de recherche afin de notifier les documents pertinents identifiés (Evans and Chi, 2010). Cette considération d’un contexte interactif est renforcée par certains travaux qui se sont intéressés aux différents facteurs d’influence de la reformulation de requête en contexte collaboratif (Yue et al., 2014). Ils ont souligné que les pages visitées et jugées comme pertinentes par un collaborateur peuvent impacter sur ses requêtes reformulées ainsi qu’augmenter la technicité ou la diversité du vocabulaire des membres du groupe. Selon un point de vue orienté vers les modèles de RIC synchrones, Foley (2008) présente une modélisation de session de recherche qui implique, comme illustré dans la Figure 3.2 :

1. Un ensemble de requêtes successives, ainsi que leurs méta-données temporelles, traduisant des reformulations du besoin en information partagé.
2. Un ensemble des pages visitées, sauvegardées ou autres jugements de pertinence, ainsi que leurs méta-données temporelles, qui traduisent les appréciations des utilisateurs par rapport à la pertinence des documents vis-à-vis du besoin en information partagé.

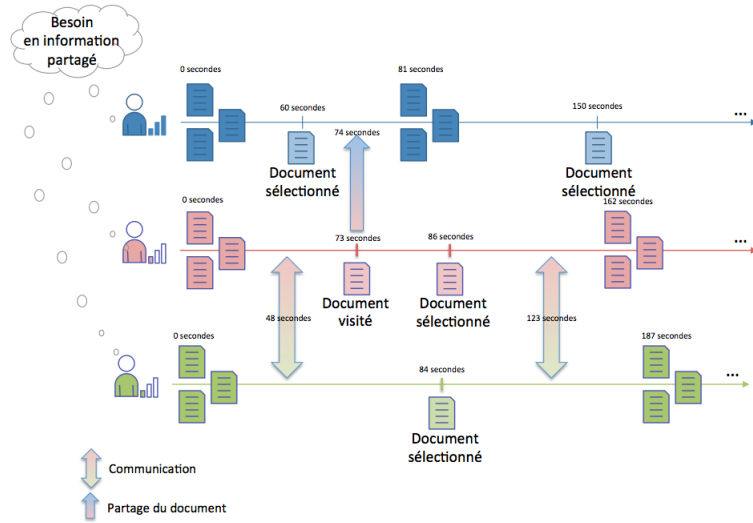


FIGURE 3.2: Exemple d'une session de RIC impliquant trois utilisateurs

3.3 Les modèles génériques de la collaboration

Les modèles génériques analysent le déroulement d'une session de RIC afin d'identifier les principales étapes qui peuvent être reliées, par exemple, à des phases du comportement utilisateur pouvant traduire leurs stratégies ou leur implication dans la session de recherche.

En RI individuelle, de nombreux modèles (Belkin, 1980; Ellis, 1989; Kuhlthau, 1991; Xie and Joo, 2012) se sont intéressés à modéliser le comportement d'un utilisateur unique. Il existe cependant des différences fondamentales entre des processus individuels et collaboratifs (Böhm et al., 2013; Joho et al., 2008; Karunakaran et al., 2013; Shah and González-Ibáñez, 2011a), telles que la diversité du vocabulaire utilisé et l'efficacité de la session. En outre, le processus de collaboration induit des difficultés supplémentaires liées au fait que la place de la communication est plus importante, le besoin en information est plus complexe et les informations collectées doivent être agrégées à l'ensemble des collaborateurs. Ces difficultés sont liées à la complexité d'une session de RIC qui comporte quatre principaux éléments (Shah, 2012) : les sources d'information, les outils technologiques de collaboration, les utilisateurs et le résultat de la collaboration.

Initié comme modèle générique de RI individuelle, le modèle ISP (Kuhlthau, 1991), illustré dans la Figure 3.3, est un des modèles de comportement le plus connu qui inclut cinq phases : initiation, sélection/exploration, formulation, collection et présentation. Hyldegård (2006, 2009) a analysé le processus de RIC afin d'identifier si le modèle ISP peut être appliqué dans ce contexte. Bien qu'il existe des similarités entre les processus individuels et collaboratifs vis-à-

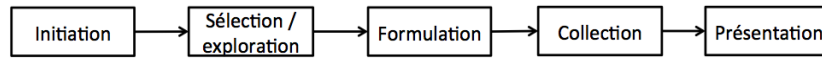


FIGURE 3.3: Etapes du modèle ISP (Kuhlthau, 1991)

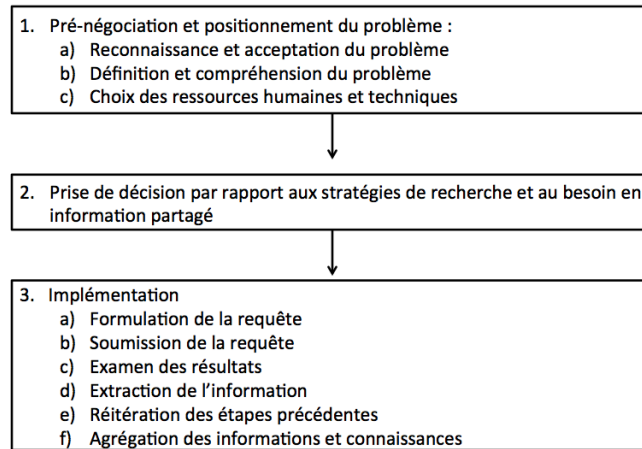


FIGURE 3.4: Etapes fusionnées des deux modèles de RIC proposés par Gray (1989) et Marchionini (1995)

vis du modèle ISP, les auteurs ont démontré que certaines dimensions sociales et cognitives doivent être ajoutées dans la mesure où elles affectent la collaboration. Ainsi, les facteurs sociaux et ceux liés à la tâche permettent une meilleure adaptation à la complexité du contexte multi-utilisateurs (Hyldegård, 2006).

Dans une autre catégorie de contributions (Gray, 1989; Evans and Chi, 2010), le processus de collaboration n'est pas dérivé des modélisations issues de la RI individuelle. Par conséquent, ces nouvelles approches permettent d'intégrer, tels des facteurs intrinsèques, les composants de la collaboration.

Une des premières contributions dans cet optique est celle proposée par Gray (1989), qui a scindé le processus de collaboration en trois phases : (1) la pré-négociation ou positionnement du problème, (2) la prise de décision et (3) l'implémentation. Shah (2012) combine ce processus avec neuf sous-scénarios proposés par Marchionini (1995), comme illustré dans la Figure 3.4.

Axés sur une approche sociale, Evans and Chi (2010) ont proposé un modèle de RI, illustré dans la Figure 3.5, qui prend en compte les interactions entre utilisateurs pour la résolution d'un besoin en information partagé. Pour cela, les interactions peuvent être réalisées par l'utilisation d'un réseau social ou d'expertise, d'un espace de travail partagé, des connaissances extraites de l'intelligence collective afin d'améliorer la recherche (Evans and Chi, 2010). Quelle que soit la nature de la tâche (informationnelle, navigationnelle ou transactionnelle), Evans

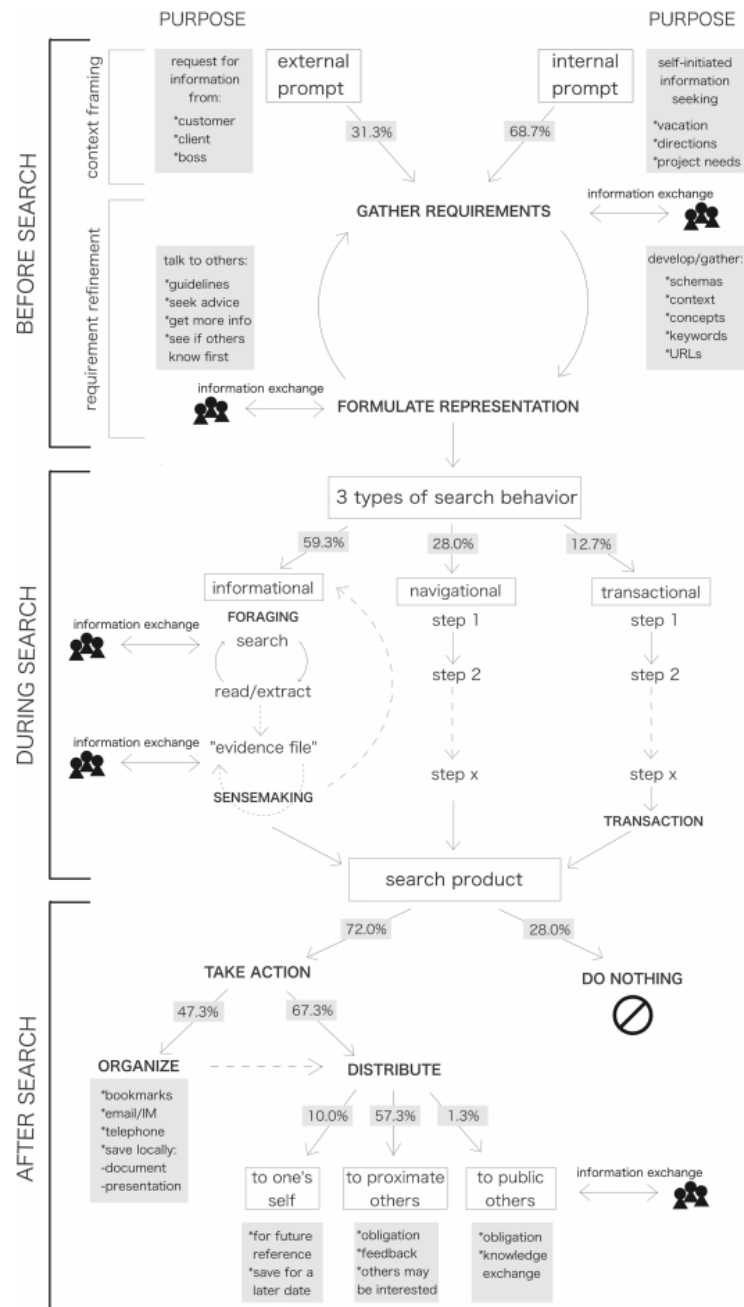


FIGURE 3.5: Modèle de RI sociale basé sur les interactions entre collaborateurs (Evans and Chi, 2010)

and Chi (2010) identifient trois étapes lors du processus de collaboration liées à la dimension temporelle :

- *Avant la recherche*, les utilisateurs échangent afin de clarifier leur besoin en information qui émane d'une source externe ou d'une motivation personnelle. La situation de collaboration peut être motivée par le besoin de conseils ou d'un retour d'expertise.
- *Pendant la recherche*, les utilisateurs interagissent afin d'établir une pertinence collective par la confrontation des opinions, des séances de réflexion collective et l'amélioration du schéma de recherche. La collaboration, prédominante pour les requêtes informationnelles, consiste à récupérer l'information et à lui donner du sens afin de reformuler éventuellement la requête.
- *Après la recherche*, les utilisateurs organisent et/ou partagent l'information récoltée. L'organisation est visible par des annotations, des bookmarks, la constitution de dossiers ou encore l'impression des résultats pour une utilisation future. Le partage des ressources est généralement effectué dans un rayon proche, amis et/ou collègues, motivé par une demande externe d'information *–expressed interest–* (Capra et al., 2010) ou la possibilité de l'intérêt des autres utilisateurs pour l'information récoltée *–Perceived interest–* (Capra et al., 2010).

Cette approche est également adoptée par Karunakaran et al. (2013) qui distinguent trois phases dans le comportement des collaborateurs au cours de la session. Comme indiqué dans la Figure 3.6, Karunakaran et al. (2013) modélisent des phases chronologiques similaires à celles proposées par Evans and Chi (2010), faisant appel aux étapes de formulation du besoin en information (*avant*), de recherche (*pendant*) et d'utilisation de l'information (*après*).

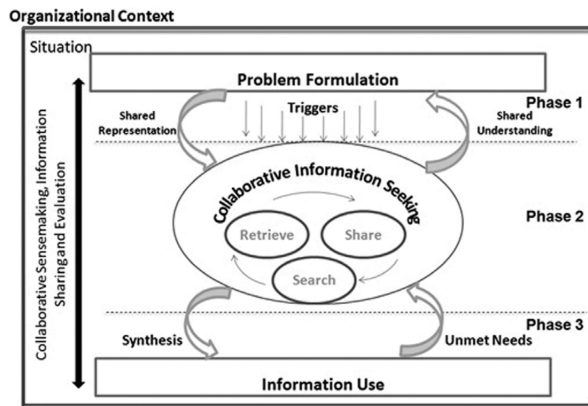


FIGURE 3.6: Modèle du comportement de l'utilisateur en contexte de RIC (Karunakaran et al., 2013)

3.4 Domaines d'application

La RI collaborative peut être appliquée dans différents domaines tels que la médecine (Sarcevic, 2009; Morris and Morris, 2011), la gestion de l'information pour les bibliothèques (Kuhlthau et al., 1992; Taylor, 1968) ou les organisations (Tungare et al., 2010; Attfield et al., 2010), la politique (Mascaro and S.P., 2010) ou le domaine académique (Hyldegård, 2006). Nous détaillons ci-après des illustrations du processus de RIC selon certains domaines d'application. Notons que cette liste est non-exhaustive et qu'il existe de nombreux domaines d'application de la RIC.

3.4.1 Domaine médical

Le principal objectif de la collaboration dans le domaine médical réside dans l'identification d'informations à propos de maladies, de prise en charge d'un patient ou d'un processus médicamenteux souvent complexe du fait de l'hétérogénéité des patients, des traitements et du grand nombre de bases de connaissances médicales (Foster, 2006; Hertzum, 2008; Morris and Morris, 2011). Un réel besoin a été exprimé dans le domaine médical envers des systèmes d'information collaboratifs, permettant l'échange entre patients, médecins et personnels soignants (Golovchinsky et al., 2009b; Morris and Morris, 2011; Hertzum, 2008). Dans ce contexte, Wald et al. (2007) introduisent la notion de triangularisation entre les patients, le web et les médecins. Cette dernière fait référence aux relations entre ces différentes entités et débouche sur des activités concrètes telles que l'aide à la prise de décision pour un patient en consultation avec son médecin à partir d'informations extraites du Web. Dans le contexte d'une collaboration patient-médecin, les patients sont des consommateurs actifs d'information médicale (ECDPC, 2011) afin de préparer une future consultation (Attfield et al., 2006; Fox, 2008) alors que les médecins sont, quant à eux, des aides dans le processus de RI. En effet, Podichetty et al. (2006) ont mis en évidence que 80% des professionnels de la santé ont déjà partagé de l'information qu'ils avaient récupéré du Web avec leur patient pendant une consultation. Ainsi, patients et médecins collaborent afin d'analyser les informations liées à la maladie du patient (McMullan, 2006). La collaboration peut également avoir lieu entre membres de l'équipe soignante (Hertzum, 2008; Lange, 1998; Reddy and Spence, 2002), qui dans un premier temps se coordonnent au quotidien pour assurer le bien-être des patients, mais également interagissent pour échanger de l'information ou confirmer des avis sur un sujet donné (Dawes and Sampson, 2003). Dans le cas d'un processus médicamenteux, le processus fait interagir plusieurs intervenants, notamment les médecins qui prescrivent le traitement et les infirmières qui, plus souvent en contact avec le patient, donnent des éléments sur les symptômes (Hertzum, 2008). La coordination et la collaboration entre ces différents acteurs est donc primordiale pour assurer les soins du patient.

3.4.2 Domaine de la gestion d'information

Ce domaine est vaste et fait référence à la compréhension des informations personnelles (Tungare et al., 2010), mais également à la gestion des connaissances, par exemple dans les bibliothèques (Taylor, 1968) ou dans les entreprises avec le domaine du e-Discovery (Attfield et al., 2010).

Domaine bibliothécaire. Dans la tâche de recherche bibliographique, un outil en ligne a été proposé : *QuestionPoint*¹. C'est un service qui permet d'interagir avec des bibliothécaires experts ayant pour objectif de suggérer des requêtes, des conseils de navigation ou également de fournir l'information pertinente. En outre, de nombreux travaux se sont intéressés à la collaboration entre un bibliothécaire et un individu où la tâche commune est de satisfaire le besoin de l'individu en s'appuyant sur l'aide du bibliothécaire (Kuhlthau et al., 1992; Taylor, 1968; Wielhorski, 1994). Cette collaboration permet aux individus de gagner du temps et de diminuer l'effort cognitif. De plus, les interactions avec le bibliothécaire permettent à l'individu de réajuster son besoin en information, mais également d'éduquer l'individu à la recherche bibliographique pour qu'il devienne autonome (Wielhorski, 1994).

Domaine du e-Discovery. Le domaine de la e-Discovery s'intéresse à la gestion des documents électroniques dans les organisations ou les entreprises dans l'objectif de les réutiliser en cas de litige civil ou criminel ou d'inspection gouvernementale (Conrad, 2007). La complexité de cette gestion réside dans le fait que malgré des compétences et qualifications élevées de la part des juristes, une collaboration entre les acteurs (entreprises, juristes et inspections gouvernementales) permet (a) de retrouver les documents *–privileged material–* qui répondent à une situation critique, (b) de discuter des points sensibles et (c) d'élaborer une conscience mutuelle (Attfield et al., 2010; Yue and He, 2010). En effet, certains travaux soulignent que le besoin en information n'est pas directement formulé par les individus qui recherchent de l'information et collaborent (Attfield et al., 2010). De plus, la pertinence des documents n'est pas limpide (Wang and Soergel, 2010). Dans ce contexte, deux rôles principaux sont identifiés (Attfield et al., 2010) : les *fee-earners* qui travaillent avec les clients et sollicitent l'information et les *knowledge management staff* qui fournissent les outils et ressources de gestion de l'information ainsi que des connaissances. Ces deux rôles collaborent avec les clients pour identifier leur besoin et fournir les documents utiles. Egalement, les juristes collaborent entre eux afin que les moins expérimentés *–contact reviewers–* bénéficient de l'expérience des experts *–lead counsel–* (Privault et al., 2010; Wang and Soergel, 2010).

1. <http://www.iclc.org/questionpoint/>

3.4.3 Domaine académique

De nombreux travaux (Foster, 2006; Hyldegård, 2006; Large et al., 2002; Amershi and Morris, 2008; Moraveji et al., 2011) se sont intéressés aux tâches de collaboration dans un contexte d'enseignement ainsi qu'à l'apport de la collaboration sur l'apprentissage. Les étudiants sont amenés à collaborer sur des tâches de RI dans leur scolarité, pour la réalisation de projets par exemple (Large et al., 2002; Amershi and Morris, 2008; Moraveji et al., 2011). L'utilisation d'un système de visualisation en temps réel de la classe, qui supporte les interactions sociales entre les élèves et avec le professeur a démontré sa valeur pédagogique dans la qualité d'apprentissage (Moraveji et al., 2011). Le professeur peut voir les élèves travailler, identifier ceux qui ont des problèmes et permettre aux élèves d'agréger leurs actions et de répondre collectivement à un problème donné par le professeur.

3.5 Approches de la RIC

Dans cette section, nous nous intéressons aux différentes approches proposées comme support à la RIC prenant en compte la problématique liée au contexte multi-utilisateurs. Nous délimitons dans un premier temps la problématique de ces approches afin de distinguer ensuite deux grandes catégories de travaux qui proposent des approches basées sur une médiation utilisateur (interfaces adaptées à la collaboration) et des solutions basées sur une médiation algorithmique reposant sur les approches liées aux modèles et techniques de RI adaptés à la collaboration.

De nombreuses problématiques sont soulevées par l'introduction du contexte multi-utilisateurs dans un processus de recherche où la collaboration est explicite :

1. Comment organiser et structurer la collaboration afin d'optimiser les actions des utilisateurs ?
2. Comment garantir l'efficacité de la session afin de résoudre le besoin en information partagé ?

Pour répondre à ces enjeux, deux principales approches, détaillées dans les sections suivantes, sont proposées dans la littérature (Joho et al., 2009; Kelly and Payne, 2013) :

1. *Les approches basées la médiation utilisateur* reposent principalement sur la définition d'interfaces adaptées à la collaboration afin de faciliter les interactions entre utilisateurs liées à l'organisation mais également les échanges liés aux résultats analysés au cours de la session. Pour cela, les interfaces disposent généralement d'outils de communication (Morris and Horvitz, 2007) ainsi que d'un espace de travail partagé (Shah and González-Ibáñez, 2011a).

2. *Les approches basées la médiation système* se focalisent sur l’ordonnancement collaboratif des documents en rapport avec le besoin en information partagé. De plus, ces approches s’intéressent particulièrement à la garantie du paradigme de division du travail en proposant de séparer les tâches des utilisateurs (Pickens et al., 2008; Shah et al., 2010) et/ou les espaces documentaires à explorer (Foley and Smeaton, 2009; Morris et al., 2008).

3.5.1 Approches basées sur la médiation utilisateur

Les interfaces collaboratives (Hartmann et al., 2009; Rodriguez Perez et al., 2011; Morris and Horvitz, 2007) reposent principalement sur une médiation utilisateur et fournissent principalement des outils de coordination et d’organisation afin de laisser les collaborateurs libres de leurs actions (Kelly and Payne, 2013). Nous présentons dans ce qui suit les principales caractéristiques des approches de médiation utilisateur ainsi que quelques exemples d’interfaces collaboratives.

3.5.1.1 Caractéristiques des approches de médiation utilisateur

La synthèse des points clés de la collaboration mise en évidence par Thomson and Perry (2006) et les conseils pour l’élaboration de systèmes de collaboration proposés par Shah (2012) nous permettent de dégager trois principaux composants des interfaces collaboratives :

1. *Un moyen de communication entre les utilisateurs.* Un espace de communication, tel une messagerie instantannée, permet aux collaborateurs d’établir les stratégies de recherche, de déterminer les rôles et les responsabilités de chacun ainsi que d’échanger sur la pertinence des documents visualisés (Golovchinsky et al., 2011; Kelly and Payne, 2013).
2. *Un espace de visualisation des actions de la communauté.* Cet espace fait appel au paradigme de conscience de l’environnement collaboratif –*awareness*– (Morris and Horvitz, 2007), permettant également aux utilisateurs d’explorer de plus larges domaines à partir des recherches des collaborateurs. En effet, Joho et al. (2008) ont mis en évidence le besoin des utilisateurs de visualiser des résumés de la collaboration afin de donner du sens à leurs actions (Paul and Morris, 2009). Pour cela, le système peut inclure un historique de recherche (Diriye and Golovchinsky, 2012), des vidéos des différentes actions (Rodriguez Perez et al., 2011), une double interface qui recense les actions de l’individu et celles de ses collaborateurs.
3. *Un espace de stockage collectif et individuel.* Le système doit garantir le principe du bénéfice mutuel de la collaboration. Une des solutions proposées est d’intégrer un espace de travail privatif et collectif.

3.5.1.2 Exemples d'interfaces collaboratives

Une première catégorie d'interfaces (Erickson, 2010; Rodriguez Perez et al., 2011; Shah and González-Ibáñez, 2011a; Vivian and Dinet, 2008) propose un outil logiciel, souvent disponible sur le web, supportant la collaboration par des techniques de visualisation ou en proposant des espaces de travail partagé. Par exemple, *Coagmento*² (Shah and González-Ibáñez, 2011a) est un système de collaboration reposant sur un plugin Firefox qui permet aux collaborateurs, tout en naviguant sur le Web, de sauvegarder des documents. Les utilisateurs peuvent alors collaborer selon différents scénarios : synchrone/asynchrone, proche/éloigné. Ce plugin est utilisé dans de nombreuses évaluations expérimentales sur l'analyse du comportement en situation de RI (González-Ibáñez et al., 2013; Shah and González-Ibáñez, 2011b), facilitées par la mise en ligne du code du plugin³.

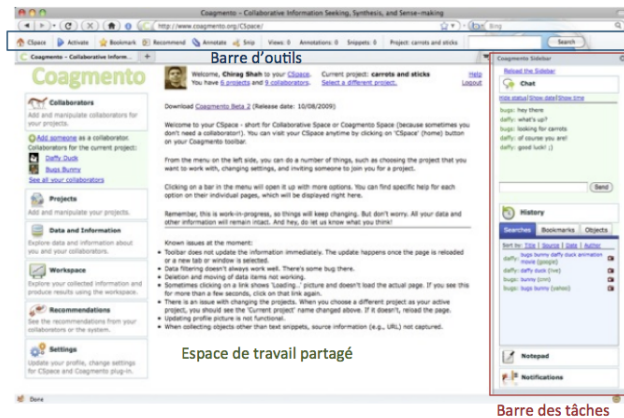


FIGURE 3.7: Plugin *Coagmento* (Shah and González-Ibáñez, 2011a)

La Figure 3.7 présente les différents composants de *Coagmento* qui intègrent les fonctionnalités suivantes :

- Un espace de travail gérant plusieurs projets où chacun d'eux peut être administré de façon collaborative ou individuelle.
- Un plugin Firefox composé (a) d'une barre d'outils permettant aux collaborateurs d'annoter, de commenter ou sauvegarder les documents ou morceaux d'information pertinents ainsi que (b) d'une barre des tâches qui inclut un système de communication entre collaborateurs et de visualisation des actions réalisées par le groupe.

2. <http://coagmento.org/>

3. <http://www.coagmento.org/collaboratory.php>

A l'inverse de la précédente interface qui stocke les actions des utilisateurs dans un espace partagé, Rodriguez Perez et al. (2011) proposent une interface de collaboration synchrone, appelée *CoFox*, qui repose sur la spécificité de résumer les actions des utilisateurs et les pages visitées dans un flux de vidéos. Les avantages de ce principe sont multiples : représentation visuelle des actions et des informations visitées, interface agréable, dynamique et facile d'utilisation. L'interface, illustrée dans la Figure 3.8, présente deux parties. Le volet de droite de l'interface est dédié à l'utilisateur local et à sa navigation tandis que le volet de gauche est dédié au collaborateur éloigné et permet à l'utilisateur local de voir ses actions, de dialoguer avec lui, de partager des liens et de regarder en temps-réel sa navigation.

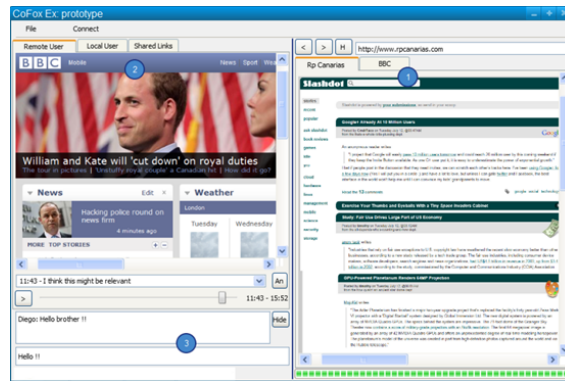


FIGURE 3.8: Interface *CoFox* (Rodriguez Perez et al., 2011)

La deuxième catégorie d'interfaces (Hartmann et al., 2009; Morris et al., 2006, 2010; Smeaton et al., 2006) reposant sur une médiation utilisateur propose des technologies plus évoluées, telles que les tables tactiles, qui s'appuient sur des espaces de travail communs, à taille humaine, permettant une interaction simultanée et en face-à-face entre utilisateurs ainsi que la coordination, conscience et visibilité des actions des collaborateurs. Par exemple, Smeaton et al. (2006) ont proposé un système de table tactile, appelé *Fischlar-DiamondTouch* dévoué à la recherche d'images. Comme illustré dans la Figure 3.9, les utilisateurs sélectionnent et organisent les documents en face-à-face, ce qui leur permet de communiquer et d'échanger sur le besoin en information.

Egalement sur la base d'une table tactile dévouée à la recherche d'images, Morris et al. (2006) ont réalisé une interface appelée *TeamSearch*, destinée à un groupe de quatre utilisateurs. Ce système, illustré dans la Figure 3.10, aide les petits groupes de collaborateurs par un système d'annotations qui permet de collecter et catégoriser les images. Au lieu de soumettre leurs requêtes dans un champ texte, les utilisateurs sélectionnent des éléments de catégories pour construire des requêtes, visualisées par un insigne de couleur différente pour chaque membre du groupe pour assurer la transparence des actions. Dans le même esprit, Morris et al. (2010) proposent une grande table de travail interactive permettant aux

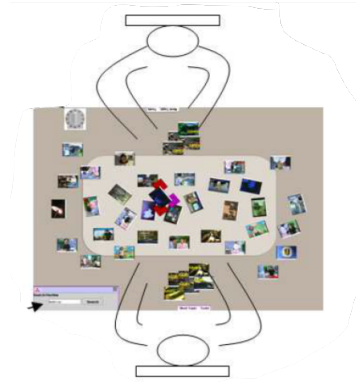


FIGURE 3.9: Système *Fischlar-DiamondTouch* (Smeaton et al., 2006)

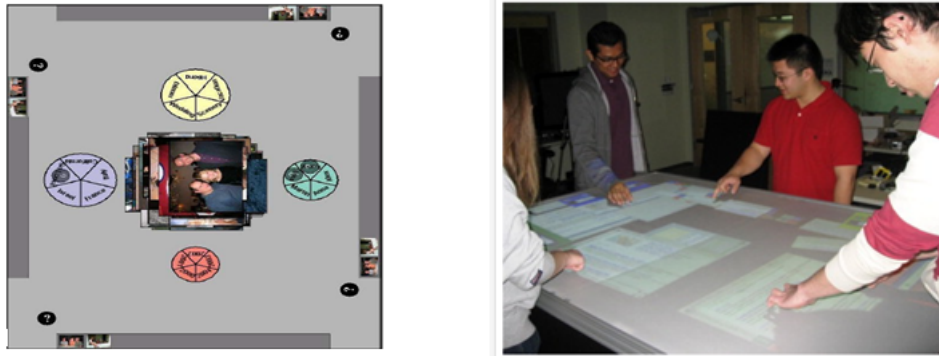


FIGURE 3.10: Système de collaboration reposant sur une table de travail interactive, respectivement TeamSearch (Morris et al., 2006) et WeSearch (Morris et al., 2010)

utilisateurs de chercher sur le web en interagissant avec les autres utilisateurs. Ce système, appelé *WeSearch*, offre la possibilité aux utilisateurs d'avoir des espaces séparés mais visibles par l'ensemble des membres du groupe pour faciliter le paradigme de la conscience de l'environnement collaboratif. En outre, *WeSearch* intègre un générateur de fichier XML contenant les résultats de recherche des utilisateurs. Ce système est illustré dans la Figure 3.10.

3.5.1.3 Synthèse des caractéristiques des interfaces présentées

Les caractéristiques de l'ensemble des systèmes présentés dans la section précédente sont synthétisées dans le tableau 3.2 en tenant compte des dimensions, à savoir le niveau de la localisation, la concurrence, du paradigme d'*awareness* et de la taille possible du groupe de collaboration. Notons que la liste des in-

terfaces est non exhaustive. Nous avons privilégié les plus connues, mais il en existe d'autres tout aussi intéressantes (Haraty et al., 2010; Hopfgartner et al., 2008; Isenberg and Fisher, 2011).

		Coagmento	CoFox	Fischlar-DiamondTouch	TeamSearch	FourBySix	SearchTogether
Localisation	éloignée	■	■	□	□	□	■
	proche	■	■	■	■	■	■
Concurrence	synchrone	■	■	■	■	■	■
	asynchrone	■	□	□	□	□	■
<i>Awareness</i>	espace de travail partagé	■	□	□	□	□	□
	affichage des actions	■	■	□	□	□	■
	tables tactiles	□	□	■	■	■	□
Taille du groupe	paire	□	□	■	□	□	□
	petit groupe (2-6)	□	□	□	■	■	□
	groupe plus large (2 et plus)	■	■	□	□	□	■

TABLE 3.2: Synthèse des approches basées sur la médiation utilisateur

3.5.2 Approches basées sur la médiation système

Les approches basées sur la médiation système permettent de supporter la collaboration d'un point de vue algorithmique afin d'ordonnancer les documents tout en garantissant les paradigmes de la collaboration. Dans cette section, nous énonçons les enjeux et les caractéristiques de ces approches de collaboration basées sur la médiation système.

3.5.2.1 Enjeux des modèles de RIC

La RIC considère un scénario de recherche caractérisé par plusieurs utilisateurs qui collaborent explicitement afin de résoudre un besoin en information commun. L'ensemble des interactions effectuées durant la session permet aux collaborateurs d'affiner leur besoin en information ainsi que leurs stratégies de recherche et peut être utilisées comme source d'évidence dans l'estimation de la pertinence des documents. L'intuition globale des approches collaboratives basées sur une médiation système est de définir une fonction d'estimation de la pertinence des documents, qui, à partir d'un besoin en information partagé, retourne un ou plusieurs ordonnancements.

Le contexte interactif lié à une tâche de RIC soulève de nombreux enjeux dans l'estimation de la pertinence des documents :

1. *La construction du profil des collaborateurs dans le processus de RIC.* L'objectif est de comprendre, à partir des jugements de pertinence des collaborateurs –*feedback*–, leur comportement afin d'identifier leur rôle dans la collaboration ou de tenir compte de leurs préférences par rapport aux documents visités ou leurs stratégies de recherche.
2. *L'estimation de la pertinence collective et individuelle des documents vis-à-vis du comportement de collaboration des utilisateurs.* L'objectif est d'estimer le score de pertinence des documents à partir du besoin en information partagé qui personnalise également les résultats par rapport au profil ou au rôle des collaborateurs.
3. *La garantie des paradigmes de la collaboration.* L'intuition est de favoriser un effet synergique entre les collaborateurs. Le principe de division du travail est celui qui est le plus utilisé car il permet de structurer et d'organiser la collaboration. La division du travail peut être déclinée selon une approche de distinction basée sur les ensembles documentaires ou les tâches des collaborateurs (Foley and Smeaton, 2010).

Pour répondre à ces enjeux, des travaux ont été proposés faisant émerger deux axes communs qui reposent sur la prise en compte des jugements de pertinence –*feedback*– des collaborateurs et de la division du travail. Une des principales distinctions réside dans les rôles des utilisateurs qui peuvent être distingués selon s'ils sont (a) symétriques, lorsque les utilisateurs sont considérés comme des pairs et ont le même objectif, ou (b) asymétriques avec des objectifs distincts. Par conséquent, nous proposons de présenter les approches selon cette catégorisation.

3.5.2.2 Modèles d'ordonnement collaboratif basés sur les rôles symétriques

L'essence de ces modèles est de considérer les utilisateurs de façon symétrique avec des rôles de pairs (Golovchinsky et al., 2009b). Dans ce contexte, la division du travail durant la tâche collaborative est assurée par une distinction

des espaces documentaires explorés par les collaborateurs afin de répondre au besoin en information partagé. Pour cela, deux principales techniques sont utilisées. La plus naïve (Foley, 2008; Foley and Smeaton, 2009) consiste à estimer la pertinence collective des documents et de les distribuer de façon indifférente aux collaborateurs tandis que la deuxième repose sur une “division intelligente” –*smart-splitting*– (Morris et al., 2008) qui personnalise les listes de résultats par rapport à des profils utilisateurs.

Foley and Smeaton (2009) proposent un modèle qui agrège les jugements de pertinence des collaborateurs afin d’ordonnancer les documents en réponse à une requête individuelle formulée par un utilisateur au cours de la session de RIC. Plus particulièrement, l’agrégation est réalisée au niveau de la pondération $purw(v)$ des termes t_v comme un poids de pertinence partiel orienté utilisateur –*partial-user relevance weighting*– :

$$purw(t_v) = \log \frac{(\sum_{u=0}^{U-1} \alpha_u \frac{r_{uv}}{R_u})(1 - \sum_{u=0}^{U-1} \alpha_u \frac{n_v - r_{uv}}{N - R_u})}{(\sum_{u=0}^{U-1} \alpha_u \frac{n_v - r_{uv}}{N - R_u})(1 - \sum_{u=0}^{U-1} \alpha_u \frac{r_{uk}}{R_u})} \quad (3.1)$$

où r_{uv} représente le nombre de documents contenant le terme t_v et jugés comme pertinents par l’utilisateur u . Le nombre de documents jugés comme pertinents par l’utilisateur u est noté R_u . Le coefficient α_u , appelé également facteur d’autorité, correspond à l’impact de l’utilisateur u dans le poids du terme, sous la contrainte que $\sum_{u=0}^{U-1} \alpha_u = 1$. Le facteur d’autorité peut être estimé 1) de façon statique, fixé à l’expérimentation du modèle ou 2) de façon dynamique en estimant la corrélation entre les poids des termes issus des documents jugés comme pertinent par l’utilisateur et les poids des termes dans un ensemble de documents appelés *oracle*, constituant l’ensemble des documents pertinents pour la requête traitée. Cet oracle est construit à partir d’hypothèses de pertinence, telles que l’ensemble des jugements de pertinence effectués par un ensemble de groupes de collaboration.

Les termes peuvent être réutilisés pour deux tâches de RI :

1. L’ordonnancement des documents qui intègre le poids de pertinence partiel des termes orienté utilisateur dans le modèle probabiliste proposé par Robertson and Walker (1994). Le score de pertinence $RSV(d_i, q_h)$ d’un document d_i en réponse à une requête q_h est alors estimé comme suit :

$$RSV(d_i, q_h) = \prod_{t_v \in q_h} purw(t_v) \quad (3.2)$$

Le principe de division du travail est ensuite assuré en ne retournant que les documents qui n’ont pas été visités par les autres collaborateurs et/ou qui ne sont pas couramment inclus dans leurs ordonnancements. Nous soulignons le fait que l’agrégation des jugements de pertinence de l’ensemble du groupe de collaboration, bien que provoquant l’absence de personnalisation des ordonnancements, assure également l’application du paradigme de partage des connaissances.

2. L'expansion de requêtes (Foley, 2008) où le poids $puow(t_k)$ –*partial-user offer weighting*– du terme t_k est estimé ainsi :

$$puow(t_k) = \sum_{u=0}^{U-1} (\alpha_u \cdot r_{uv} \cdot purw(t_k)) \quad (3.3)$$

Ce modèle a été évalué par une approche de simulation de la collaboration à partir de fichiers log de la collection TREC Interactive. Après avoir effectué un paramétrage sur le facteur d'autorité qui obtient une valeur optimale lorsqu'il est calculé de façon dynamique, les principales analyses mettent en évidence qu'une division du travail totale (ordonnancements distincts et suppression des documents visités) assure une efficacité et une diversité plus importante qu'une division du travail partielle. De plus, les résultats montrent que l'efficacité du modèle collaboratif est supérieure à celle du meilleur utilisateur participant à la collaboration. Bien que ce modèle présente l'avantage d'estimer la pertinence collective des documents en réponse à un besoin en information partagé, il occulte cependant l'aspect personnalisation qui est nécessaire à la satisfaction du bénéfice mutuel.

Dans cette optique de personnalisation, Morris et al. (2008) proposent deux modèles collaboratifs basés sur le score de personnalisation $perso(d_i, u_j)$ d'un document d_i par rapport à un utilisateur u_j , défini par Teevan et al. (2005) et présenté dans l'Equation 2.17 de la section 2.2.1 :

1. Le premier modèle, appelé *smart-splitting*, permet de distribuer, à partir d'une seule requête utilisateur, les documents d_i entre les collaborateurs par une méthode de personnalisation. Dans cet objectif, les auteurs assignent chaque document d_i au collaborateur u_j qui obtient le score de personnalisation $perso(d_i, q_h, u_j)$ le plus élevé, comme détaillé ci-dessous :

$$\forall d_i, \exists u_j^*; u_j^* = \arg \max_{u_j \in U} perso(d_i, q_h, u_j) \quad (3.4)$$

2. Le deuxième modèle, appelé *groupization*, a été proposé sur la base du même score personnalisé $perso(d_i, q_h, u_j)$ (Teevan et al., 2005) en les agrégeant afin d'ordonner les documents d_i selon une notion de pertinence collective. Le score de groupization $g(d_i)$ est ensuite combiné avec le rang $rang(d_i)$ original du document d_i pour un moteur de recherche web classique afin de préserver les informations les plus importantes. Ainsi, le score collaboratif $s_{coll}(d_i)$ d'un document d_i est estimé comme suit :

$$s_{coll}(d_i) = \alpha \sum_{u_j \in U} perso(d_i, q_h, u_j) + (1 - \alpha)rang(d_i) \quad (3.5)$$

Les expérimentations réalisées pour valider ces deux modèles ont considéré des fichiers log d'une expérimentation réelle de tâche collaborative. Les résultats montrent l'effet synergique des deux modèles ("*smart-splitting*" et "*groupization*"), considérés de façon indépendante, par rapport à une tâche individuelle de recherche sur le web. Plus particulièrement, la fonction d'ordonnement

“*smart-splitting*” permet une mise en place de la division du travail plus efficace qu’une simple séparation des ensembles documentaires aléatoire. Notons que ces modèles ont été intégrés dans des interfaces de médiation hybride utilisateur-système (Morris and Horvitz, 2007).

3.5.2.3 Modèles d’ordonnancement collaboratif basés sur les rôles asymétriques

Les approches présentées jusqu’alors considèrent les utilisateurs comme des pairs, effectuant la tâche symétrique d’identifier le maximum de documents pertinents en rapport avec un besoin en information partagé. Cependant, d’autres travaux (Pickens et al., 2008; Shah et al., 2010) ont proposé d’intégrer les rôles des utilisateurs dans le processus de collaboration afin de structurer et d’organiser les actions des membres du groupe (Kelly and Payne, 2013). A l’inverse des précédents modèles où le nombre de collaborateurs est illimité, les modèles présentés ci-dessous se basent sur une paire de rôles, restreignant par conséquent la collaboration à une paire d’utilisateurs.

Pickens et al. (2008) proposent un modèle d’ordonnancement de documents basé sur les rôles asymétriques de *prospecteur* et de *mineur*, définis dans la taxonomie de Golovchinsky et al. (2009b). Le premier rôle, *prospecteur*, favorise la diversité dans les résultats en ouvrant de nouveaux champs de l’espace d’information tandis que le deuxième rôle, *mineur*, assure la qualité et la richesse des résultats explorés. Pour modéliser les rôles, le modèle tient compte des actions par rapport aux ordonnancements précédemment retournés. Ainsi, les auteurs identifient deux facteurs estimés à partir d’une liste L_h :

- Le facteur de pertinence $w_p(L_h)$ qui mesure le ratio du nombre de documents pertinents dans la liste L_h de documents retournés pour une requête q_h , noté $|pert \in L_h|$ par rapport au nombre de documents non pertinents dans cette même liste, noté $|nonpert \in L_h|$. Ce facteur est défini comme suit :

$$w_p(L_h) = \frac{|pert \in L_h|}{|nonpert \in L_h|} \quad (3.6)$$

- Le facteur de fraîcheur $w_f(L_h)$ qui mesure le ratio du nombre de documents non visités dans la liste L_h , noté $|nonvisit \in L_h|$, par rapport au nombre de documents visités dans cette même liste, noté $|visit \in L_h|$:

$$w_f(L_h) = \frac{|nonvisit \in L_h|}{|visit \in L_h|} \quad (3.7)$$

Notons que ces deux variables sont combinables et s’opposent dans le sens où pour une requête, si la valeur de la variable de pertinence est élevée, cela signifie qu’un ratio élevé de documents, par rapport au nombre de documents retournés, a été visualisé et jugé pertinents. Autrement dit, une valeur de la variable de pertinence implique généralement une valeur de la variable de fraîcheur faible.

L'intuition de la médiation est la suivante : plus il y a de documents pertinents dans la liste et moins de documents ont été visités, plus la piste d'exploration est bonne favorisant ainsi la diversité et la pertinence du domaine par rapport à la requête.

En fonction des objectifs de leur rôle, le modèle propose deux fonctions d'ordonnement :

1. La fonction de médiation pour le rôle du *prospecteur* consiste en une suggestion de termes pour la formulation de la requête, dans le but de favoriser la diversité des résultats explorés. Ainsi, pour chaque terme t_k issu de l'ensemble des documents présents dans les listes L précédemment retournées, son score est estimé ainsi :

$$score(t_k) = \sum_{L_h \in L} w_p(L_h)w_f(L_h)rlf(t_k, L_h) \quad (3.8)$$

où $rlf(t_k, L_h)$ représente le nombre de documents de L_h qui recensent t_k .

2. Le rôle de *mineur* impose quant à lui, aux collaborateurs de regarder plus en profondeur les documents et d'identifier ceux qui répondent le plus au besoin en information partagé. Ainsi, les documents non examinés par le *prospecteur* sont réordonnés pour être proposés au *mineur* selon le score de similarité d'un document d_i estimé de la façon suivante :

$$score(d_i) = \sum_{L_h \in L} w_p(L_h)w_f(L_h)borda(d_i, L_h) \quad (3.9)$$

où la fonction $borda(d_i, L_h)$ reflétant un système de vote des documents dans une liste.

Deux prototypes, *Cerchiamo* (Adcock et al., 2007; Golovchinsky et al., 2008) et *Querium* (Diriye and Golovchinsky, 2012), on intégré ce modèle basé sur les rôles de *prospecteur* et de *mineur*. Sur la même base de fonctionnalités (soumission de requêtes, appréciation et commentaires des documents), *Cerchiamo* et *Querium* se distinguent de par la nature des informations recherchées. *Cerchiamo* consiste en une tâche d'identification de vidéos tandis que *Querium* a pour objectif de restituer des listes de documents.

Les expérimentations ont été réalisées à partir d'une expérimentation utilisateurs sur la base du prototype *Querium* reposant sur des sessions de RIC dont le besoin en information est relié à des sujets de la collection TREC Interactive. Les participants ont pu alors analyser des documents de cette même collection qui ont été ordonnés grâce au modèle collaboratif basé sur les rôles de *prospecteur* et de *mineur*. Comme les modèles précédents, les résultats assent de l'effet synergique du modèle par rapport à un scénario de recherche individuel. Une analyse au niveau des sujets, catégorisés selon deux niveaux de difficulté basés sur le nombre de documents pertinents constituant la vérité de terrain, permet de mettre en évidence que l'effet synergique du modèle est particulièrement visible lorsque les sujets sont difficiles. L'explication derrière

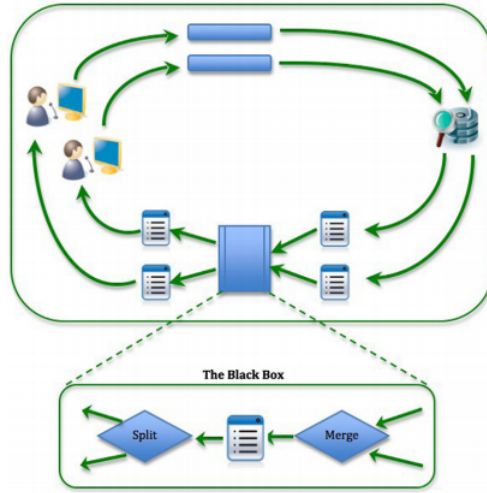


FIGURE 3.11: Principe de l'algorithme basé sur les rôles de *rassembleur* et de *surveilleur* (Shah et al., 2010)

ce constat est que si un sujet comporte beaucoup de documents pertinents, il est plus facile de les identifier et dans ce cas, une session individuelle suffit, alors que ce n'est pas le cas lorsque la tâche est plus complexe.

A partir d'autres hypothèses en terme de division des tâches, Shah et al. (2010) proposent une méthode qui s'appuie sur la paire de rôles suivante :

1. Le rôle de *rassembleur* –*gatherer*– dont le but est de détecter rapidement les documents pertinents.
2. Le rôle de *surveilleur* –*surveyor*– dont le but est de couvrir une large diversité des résultats afin de mieux comprendre la nature de la collection pour découvrir les champs exploratoires potentiels et de détecter pourquoi les requêtes ne sont pas optimales.

La collaboration repose sur la complémentarité de ces rôles. En effet, le *rassembleur* peut rechercher l'information pertinente tout seul, mais le *surveilleur* a besoin de l'intelligence collective pour obtenir une diversité thématique et documentaire pour une requête. Le principe de médiation entre ces deux rôles est illustré dans la Figure 3.11. Il repose sur la fusion et la séparation des résultats de requêtes des deux collaborateurs. Plus particulièrement, l'algorithme est déclenché à la suite des soumissions de requêtes des deux utilisateurs qui collaborent sur la base d'un même besoin en information. Pour chaque requête q_h soumise, une liste l_h ordonnée de résultats est présentée à l'utilisateur u_j associé, ainsi les listes l_h et $l_{h'}$ des deux collaborateurs u_j et $u_{j'}$ sont traitées en deux étapes :

1. Etape *Merge* où une liste ordonnée est générée par la fusion des deux ordon-

nancements grâce à la fonction CombSUM qui normalise le score $RSV(d_i, q_h)$ pour chaque document d_i dans chaque liste l_h et fait ensuite la somme des scores normalisés pour un document. L'intuition de la fusion est que la combinaison des deux ordonnancements augmente la précision de la liste fusionnée dans la mesure où elle agrège les scores $RSV(d_i, q_h)$ et $RSV(q_{h'}, d_i)$ des documents d_i les plus pertinents pour les deux requêtes q_h et $q_{h'}$. Une liste ordonnée l_{fus} est alors obtenue grâce au calcul du score de pertinence $RSV(d_i, q_h, q_{h'})$ fusionné d'un document d_i à partir des deux requêtes q_h et $q_{h'}$, estimé ainsi :

$$RSV(d_i, q_h, q_{h'}) = CombSUM(d_i, q_h, q_{h'}) = RSV_{norm}(q_h, d_i) + RSV_{norm}(q_{h'}, d_i) \quad (3.10)$$

2. Etape *Split* où un algorithme de classification, à savoir 2-means, est appliqué à la liste fusionnée l_{fus} sur la base des scores fusionnés $RSV(d_i, q_h, q_{h'})$ des documents d_i . On obtient ainsi deux ordonnancements l_{rass} et l_{surv} , respectivement retournés au *rassembleur* et au *surveilleur* selon les critères suivants :
 - (a) La classe dont le centre de gravité est le plus élevé permet de constituer la liste l_{rass} retournée au *rassembleur*.
 - (b) La classe dont le centre de gravité est le plus faible permet de constituer la liste l_{surv} retournée au *surveilleur*.

Les auteurs ont utilisé la collection TREC ad-hoc ainsi que les sujets TREC 301-450 pour évaluer leur modèle. Les expérimentations consistent principalement en un paramétrage du modèle pour déterminer les fonctions de fusion et de séparation optimales, à savoir CombSUM et l'algorithme de classification k-means. Les analyses comparent également le modèle avec un scénario de recherche individuel afin de valider son effet synergique.

3.5.2.4 Synthèse des modèles de RIC

Le tableau 3.3 résume les caractéristiques des différents modèles de RIC abordés dans cette section. Les modèles sont comparés selon leur catégorie, les ressources ordonnées, les facteurs utilisés dans la fonction d'ordonnement et les techniques de RI. Nous présentons également les avantages et inconvénients de chacun de ces modèles.

	Ordonnancement		Ressource		Facteurs		Méthode		Avantages		Inconvénients
	Interactif (durant la session)	Aggrégé (en fin de session)	Documents	Termes	Jugements de pertinence	Rôles	Autres facteurs	Probabilités	Fréquences	Fusion et redistribution	
(Morris et al., 2008) "Smart-splitting"	■	<input type="checkbox"/>	■	<input type="checkbox"/>	■	<input type="checkbox"/>	<input type="checkbox"/>	■	<input type="checkbox"/>	<input type="checkbox"/>	requête unique pour le besoin en information partagé
(Morris et al., 2008) "Groupization"	<input type="checkbox"/>	■	■	<input type="checkbox"/>	■	<input type="checkbox"/>	<input type="checkbox"/>	■	<input type="checkbox"/>	<input type="checkbox"/>	aucune personnalisation
(Foley and Smeaton, 2009)	■	<input type="checkbox"/>	■	<input type="checkbox"/>	■	<input type="checkbox"/>	<input type="checkbox"/>	■	<input type="checkbox"/>	<input type="checkbox"/>	aucune personnalisation
(Pickens et al., 2008) Prospecteur-Mineur	■	<input type="checkbox"/>	■	■	■	■	■	<input type="checkbox"/>	■	<input type="checkbox"/>	compétences de l'utilisateur pas toujours en adéquation avec les rôles
(Shah et al., 2010) Rassembleur-Surveilleur	■	<input type="checkbox"/>	■	<input type="checkbox"/>	■	■	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	■	compétences de l'utilisateur par toujours en adéquation avec les rôles

TABLE 3.3: Synthèse des caractéristiques des modèles de RIC

3.6 Evaluation de la RIC

L'évaluation de la RIC est un enjeu majeur dans la mesure où l'objectif n'est plus limité à l'évaluation de la pertinence d'un document par rapport à une requête mais plutôt à la pertinence collective reliée aux besoins de l'ensemble des utilisateurs (Shah, 2012). Dans le chapitre précédent, nous avons présenté le contexte d'évaluation dans un processus de RI individuelle où la dimension utilisateur peut être considérée comme une entité indépendante intégrée au processus de recherche. Cependant, les nouveaux enjeux liés aux modèles de RIC conduisent les travaux du domaine à adapter les protocoles d'évaluation afin de considérer (a) la dimension utilisateur par le biais de scénarios de recherche interactifs, tels que ceux basés sur la simulation ou les expérimentations utilisateurs, comme détaillé dans la section 2.1.3, et (b) la dimension de la collaboration dont l'objectif de mesurer son efficacité et son effet synergique sur les résultats de recherche. Ce dernier aspect est particulièrement traité dans cette section.

3.6.1 Problématique de l'évaluation en RIC

L'évaluation des modèles de RIC soulève de nombreux enjeux liés au contexte interactifs et collaboratifs des sessions de recherche. Dans ce qui suit, nous présentons les principaux verrous liés à l'évaluation en RIC.

La prise en compte des interactions dans le processus d'évaluation. A l'inverse de la RI individuelle où l'efficacité d'un modèle est évaluée par rapport à l'ordonnancement des documents retournés par rapport à une requête, la RIC impose la prise en compte d'une session de recherche interactive du fait du contexte multi-utilisateurs. Les interactions sont alors hétérogènes : collaborateur-document et/ou collaborateur-collaborateur. Les scénarios d'évaluation doivent alors s'adapter à cet aspect interactif afin de pouvoir évaluer des modèles collaboratifs. Pour cela, les travaux du domaine se sont basés sur la mise en place des protocoles utilisés pour la RI orientée utilisateur, tels que les *user studies* ou *log studies*.

L'estimation de la pertinence collective. A l'opposé de la RI individuelle où le besoin en information émane d'un utilisateur unique, la tâche de RIC est motivée par un besoin en information partagé par un ensemble d'utilisateurs. Ainsi, l'estimation de la pertinence est considérée comme collective afin de répondre aux attentes de l'ensemble des membres du groupe. Il s'agit d'évaluer des ordonnancements qui répondent à ce besoin partagé alors que les requêtes de la session sont reformulées et soumises de façon indépendante par différents collaborateurs.

La mesure de l'efficacité liée à la collaboration. À l'inverse de la RI individuelle où l'évaluation considère les utilisateurs comme indépendants, la RIC aborde les utilisateurs comme membres d'un groupe qui doit optimiser ses actions afin de garantir les paradigmes de la collaboration et d'atteindre un effet synergique. Plus qu'une simple agrégation des performances individuelles (Baeza-Yates and Pino, 1997), il est question de définir des mesures d'efficacité qui permettent d'évaluer, conformément au paradigme de la division du travail, l'effet synergique des sessions de RIC en vérifiant que la redondance au sein des ordonnancements soit évitée.

La mesure de l'effort cognitive de la tâche de RIC. En RI individuelle, la charge cognitive est liée seulement à la tâche de recherche consistant à résoudre un besoin en information. Dans ce contexte, la dimension de la collaboration au sein d'un processus de RIC et, par conséquent, la multiplication des interactions augmente la charge cognitive. Il existe donc un réel problème lié à la collaboration qui consiste à réduire la charge cognitive, sans dégrader la pertinence des résultats. Ainsi, il est important d'évaluer ces aspects cognitifs lors de l'évaluation d'un modèle de RIC afin de garantir l'engagement de l'utilisateur dans la session de RIC.

Notons brièvement que dans le cas d'une approche basée sur la médiation utilisateur, un des enjeux consiste en l'évaluation des interfaces de collaboration (Smyth et al., 2005; Aneiros and Morris, 2003), que nous ne détaillons pas dans ce chapitre. Brièvement, les objectifs de l'évaluation d'interface reposent sur la mesure de (a) l'efficacité des sessions de RIC réalisées au moyen de l'interface, (b) la capacité du système à faciliter la collaboration et (c) la satisfaction de l'utilisateur et son implication dans le processus de collaboration.

3.6.2 Principales approches d'évaluation

Afin de prendre en compte les interactions des utilisateurs au cours d'une session de RIC, trois protocoles, précédemment définis dans le contexte de la RI individuelle, peuvent être utilisés afin d'évaluer des modèles de RIC. Une première approche consiste à simuler la collaboration à partir de fichiers log d'utilisateurs qui ont réalisé une tâche individuelle, tandis que les deux autres approches utilisent un contexte de collaboration réel, impliquant des utilisateurs qui ont réalisé la tâche collaborative. Trois types d'évaluation sont alors possibles : (a) les expérimentations basées sur la simulation permettant de créer un contexte de collaboration, (b) les expérimentations basées sur les logs –*log study*– où les modèles de RIC sont appliqués a posteriori des actions réelles des utilisateurs et (c) les expérimentations utilisateurs –*user study*– où les participants ont interagi avec le modèle de RIC.

3.6.2.1 L'évaluation basée sur la simulation

Les premiers travaux dans le domaine ont fait face à l'absence de collections de test dédiées à la tâche de RIC. La simulation de la collaboration est une réponse à ce problème, dans la mesure où elle permet, à partir d'utilisateurs individuels, de construire une session de recherche collaboration. L'idée principale est de combiner de façon chronologique les actions (requêtes, jugements de pertinence, pages visitées, etc...) d'un ensemble d'utilisateurs afin de pouvoir simuler une session de RIC. Les modèles de RIC peuvent alors être générés à partir des requêtes de la session en prenant en compte les interactions des utilisateurs si les particularités du modèle le nécessitent. Dans leurs travaux, (Foley, 2008; Foley and Smeaton, 2009) ont utilisé cette approche d'évaluation en utilisant les fichiers log fournis par la collection TREC Interactive. La Figure 3.12 illustre un exemple de fichier log annoté qui comporte l'ensemble des requêtes soumises et des documents sélectionnés par l'utilisateur S1 qui a résolu le besoin en information 326i.

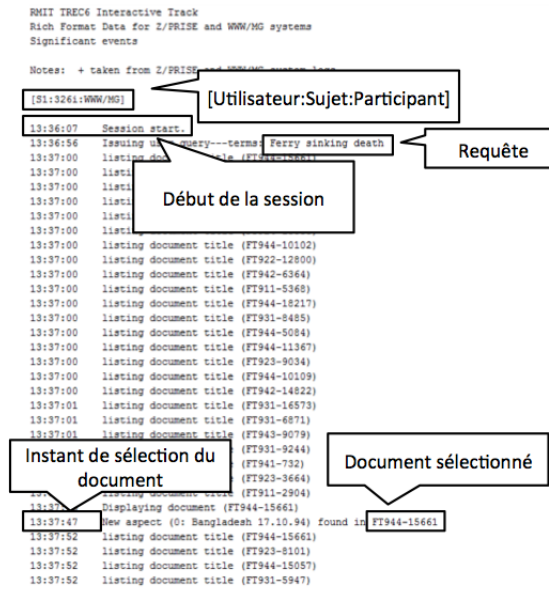


FIGURE 3.12: Exemple de fichier log fourni par le participant RmitMG6 pour le sujet TREC 326i

A partir des fichiers log, les auteurs ont constitué des paires d'utilisateurs afin de simuler une session de collaboration par la synchronisation de leur session de recherche individuelle. Compte tenu que les sessions de recherche ont été effectuées dans des intervalles de temps différents, chaque interaction est annotée par un instant qui correspond à la durée, en secondes, qui s'écoule entre le début de la session et l'action réalisée. La fusion des deux sessions individuelles

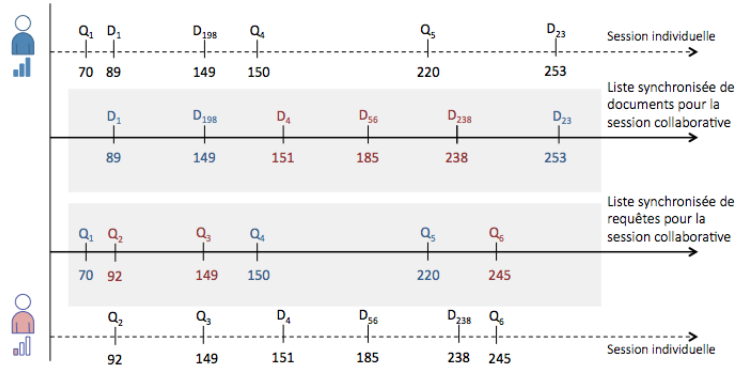


FIGURE 3.13: Simulation d'une session collaborative à partir de logs de recherche d'une session interactive individuelle

est alors réalisée de façon chronologique sur ces indicateurs temporels, comme illustré dans la Figure 3.13.

L'évaluation proposée par Morris et al. (2008) et Shah et al. (2010) est une autre forme de simulation de la collaboration selon un niveau d'abstraction plus élevé dans la mesure où aucun fichier log n'est considéré. Pour simuler la session de collaboration, Shah et al. (2010) ont demandé à 15 utilisateurs de lister les requêtes qu'ils auraient posées à chacun des 10 sujets TREC considérés. Nous soulignons le fait que les paires d'utilisateurs sont alors construites a posteriori de façon aléatoire. Le défaut de ce protocole est qu'il ne prend pas en compte les interactions des utilisateurs et ne leur donnent pas la possibilité d'affiner le besoin en information en accordance au principe de *sensemaking* relié à la collaboration. Une approche similaire, mais plus proche du contexte collaboratif, a été employée par Morris et al. (2008) en constituant 10 groupes de 3 participants qui ont eu pour objectif de lister 6 requêtes en rapport avec le besoin en information partagé.

3.6.2.2 L'évaluation basée sur les logs des utilisateurs

A l'inverse des approches qui reposent sur une simulation de la collaboration, l'évaluation basée sur les logs des utilisateurs *—log study—* considère des fichiers log de sessions de recherche collaborative réelles. L'avantage de cette approche est qu'elle considère le comportement des utilisateurs en contexte collaboratif qui est influencé par les interactions entre collaborateurs qui ont permis d'affiner le besoin en information ou les stratégies de recherche. Alors que ce type d'évaluation a été utilisé dans le cadre de la RI individuelle (Anick, 2003), il n'existe pas de travaux dans la littérature qui ont utilisé ce protocole dans un contexte collaboratif.

3.6.2.3 L'évaluation basée sur les expérimentations utilisateurs

Une autre approche consiste à considérer une tâche en temps réel où les utilisateurs interagissent avec un système collaboratif qui supporte le modèle de RIC à évaluer. Ce type d'expérimentation est appelé "*user study*" et a été mis en place dans certains travaux de la littérature (Pickens et al., 2008; Shah and González-Ibáñez, 2011b). Ces expérimentations utilisateurs reposent sur un environnement contrôlé qui implique un ensemble de participants dans une tâche supervisée (ou semi-supervisée selon les besoins de l'évaluation) où l'ensemble de leurs actions ou données sont sauvegardées.

D'une part, l'évaluation expérimentale proposée par Pickens et al. (2008) considère une tâche de recherche effectuée par des participants réels sur des collections de test, à savoir la collection TREC Interactive, où les besoins en information sont définis dans la collection de test et le modèle de RIC ordonne les documents de cette même collection. Les groupes de collaboration ont été construits à partir de 8 participants en réalisant l'ensemble des combinaisons duales possibles dans l'objectif de limiter l'effet prédominant de certains comportements individuels. D'autre part, Shah and González-Ibáñez (2011b) réalisent une expérimentation utilisateurs impliquant 10 participants individuels ainsi que 15 groupes de collaboration. Dans ce dernier travail, l'évaluation est plutôt liée à une comparaison comportementale entre les deux configurations de participants plutôt qu'à un modèle d'ordonnancement collaboratif des documents. Ce type d'évaluation permet aussi d'identifier des aspects plus subjectifs en soumettant des questionnaires aux participants (Aneiros and Morris, 2003; Morris, 2013).

3.6.3 Mesures d'évaluation

La principale limite soulevée à propos des mesures d'évaluation orientées vers un contexte de recherche individuel est qu'elles ne prennent pas en compte la dynamique de la session de recherche ainsi que la collaboration entre utilisateurs qui peuvent produire des résultats agrégés supérieurs à la somme des résultats individuels, appelé également effet synergique. En effet, ces mesures, notamment le rappel et la précision, analysent l'efficacité d'un ordonnancement indépendamment du processus de recherche réalisé. En contexte collaboratif, la pertinence des documents peut être influencée par :

1. Les actions des utilisateurs dans la mesure où les modèles de RIC (Foley and Smeaton, 2009; Pickens et al., 2008) dépendent des jugements de pertinence des utilisateurs.
2. La progression de la recherche dans le temps dans la mesure où au fur et à mesure des interactions, les utilisateurs deviennent plus familiers par rapport au besoin en information partagé (Foley, 2008).
3. Le degré de distinction de l'individu du collectif afin d'identifier la pertinence collective des documents en tenant compte du profil de chacun des collaborateurs (Foley, 2008).

Pour une meilleure complétude d'analyse des modèles de RIC, Shah (2012) propose de diviser les métriques selon deux aspects : 1) les mesures orientées système qui estiment la pertinence en terme d'efficacité d'un modèle à retourner des documents pertinents et 2) les mesures orientées individu qui s'intéressent à l'aspect cognitif et comportemental des collaborateurs. Nous détaillons ces deux types de mesure dans ce qui suit.

3.6.3.1 Mesures d'évaluation orientées système

Plusieurs métriques orientées système ont été proposées pour estimer l'efficacité d'un système de RIC (Pickens et al., 2008; Shah, 2012). Certaines d'entre elles sont inspirées de la RI individuelle, comme les mesures de rappel et précision, mais ont été adaptées au facteur collaboratif en faisant appel à la notion de couverture (Korhage, 1997). A ce jour, les mesures d'évaluation utilisées dans le domaine de la RIC ont principalement comme objectif d'estimer l'efficacité de la session en termes de résultats de recherche à partir des pages sauvegardées par les utilisateurs au cours de la session. Excepté une mesure proposée par Foley (2008), elles ne permettent donc pas d'estimer l'efficacité des ordonnancements collaboratifs de documents. L'ensemble de ces mesures est illustré dans la Figure 3.14 et est présenté ci-dessous.

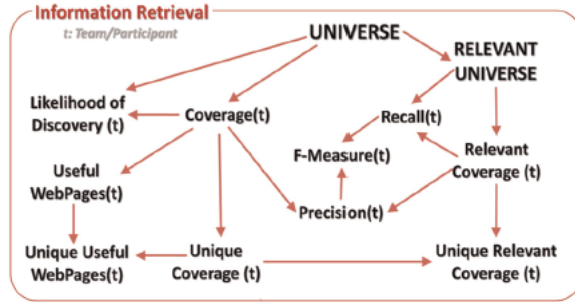


FIGURE 3.14: Mesures d'évaluation pour la RIC (Shah, 2012)

La notion de couverture. À l'inverse de certaines approches (Baeza-Yates and Pino, 1997) qui somment les résultats individuels, la couverture est une notion plus adaptée à la collaboration dans la mesure où elle permet d'agréger les actions des utilisateurs sans considérer la duplication des résultats. Ainsi, on définit la couverture $Couverture(g)$ d'un groupe g de collaboration comme l'ensemble des documents distincts visités par les membres d'un groupe g :

$$Couverture(g) = \{d_i \text{ tel que } d_i \text{ a été visité par } g\} \quad (3.11)$$

Egalement, la mesure de couverture pertinente d'un groupe g , notée $CouvPert(g)$, s'intéresse à une région particulière de l'ensemble de couverture :

$$CouvPert(g) = Couverture(g) \cap U_r \quad (3.12)$$

où U_r représente l'ensemble des documents pertinents, il est aussi appelé l'univers des pages pertinentes. Par convention, on note U l'univers des documents, c'est-à-dire l'ensemble des documents de la collection.

Nous soulignons que cette notion de couverture pertinente a été également abordée à un instant t_l de la session afin d'estimer la pertinence ponctuelle des ordonnancements retournés aux collaborateurs (Foley, 2008) en identifiant le nombre de documents distincts inclus dans les ordonnancements retournés aux collaborateurs à un instant t_l .

La notion de couverture peut également être abordée par rapport aux autres groupes de collaboration afin d'identifier les différences en terme d'efficacité par rapport à l'ensemble G des groupes impliqués dans l'évaluation. Nous pouvons donc modifier les deux mesures énoncées précédemment de la façon suivante :

$$CouvertureUnique(g) = Couverture(g) \setminus \bigcup_{g' \in G \setminus g} Couverture(g') \quad (3.13)$$

$$CouvertureUniquePertinente(g) = CouvertureUnique(g) \cap U_r \quad (3.14)$$

où $x \setminus y$ désigne le sous-ensemble de x qui ne contient pas l'ensemble y .

Une des principales limites de ces mesures est qu'elles fournissent simplement un comptage des documents sans aucune normalisation par rapport à la longueur de la session ou aux actions des utilisateurs. Il est donc difficile de les utiliser afin de comparer deux sessions de RIC. Les mesures orientées rappel et précision, décrites ci-dessous, permettent de résoudre cette faiblesse.

La précision, le rappel et la F-mesure. Comme énoncé précédemment, les mesures classiques de RI ont été adaptées à la collaboration en utilisant les notions de couverture. Ainsi, la précision, le rappel et la F-mesure, utilisées par Shah and González-Ibáñez (2011b), sont estimés ainsi :

$$Précision(g) = \frac{|CouvPert(g)|}{|Couverture(g)|} \quad (3.15)$$

$$Rappel(g) = \frac{|CouvPert(g)|}{|U_r|} \quad (3.16)$$

$$F - mesure(g) = \frac{2 \cdot Précision(g) \cdot Rappel(g)}{Précision(g) + Rappel(g)} \quad (3.17)$$

Des variantes de ces mesures ont été proposées par Pickens et al. (2008) sur la base des documents visités –*visited precision*–, des documents sélectionnés –*selected precision* et *selected recall*–.

Une des limites de ces mesures est qu'elles considèrent les documents de façon indifférenciée, sans aucune granularité dans la pertinence de la page ou, plus

proche du contexte collaboratif qui souhaite résoudre un besoin complexe, son degré de difficulté pour y accéder ou l'identifier. Des mesures implicites, basées sur les actions des utilisateurs, peuvent répondre à ce problème.

Mesures implicites Des mesures plus implicites ont été également définies en se basant sur l'hypothèse que si une page est visitée plus de trente secondes, elle doit être intéressante et utile pour l'utilisateur (Fox et al., 2005; White et al., 2010). Ainsi, Shah and González-Ibáñez (2011b) définissent la vraisemblance de découverte d'un document d_i à partir de son utilité en dérivant la mesure de l'IDF, en posant l'hypothèse que si un document est facile à trouver, sa vraisemblance de découverte est faible :

$$VraisDecouverte(d_i) = -1 \cdot \frac{|g_{d_i}|}{|G|} \quad (3.18)$$

où $|g_{d_i}|$ dénote le nombre de groupes de collaboration qui ont visité le document d_i et $|G|$ représente le nombre de groupes de collaboration impliqués dans l'expérimentation.

Illustration 3.1.

Soit une collection de 20 documents $\mathcal{D} = \{d_i ; i = 1, \dots, 20\}$. Les documents pertinents U_r –*ground truth*– pour un besoin en information \mathcal{I} sont les suivants $U_r = \{d_2, d_6, d_{15}\}$. Imaginons une session de recherche où un groupe g de collaborateurs u_1 et u_2 ont interagi avec un SRI collaboratif. Durant la session, ils ont respectivement sélectionné les documents $\{d_1, d_2, d_6, d_9, d_{17}, d_{20}\}$ et $\{d_3, d_4, d_5, d_6, d_7\}$. Nous souhaitons évaluer l'efficacité du modèle à partir du résultat de la collaboration, soit les documents sélectionnés. Les mesures d'évaluation sont estimées comme suit :

- $Couverture(g) = \{d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_9, d_{17}, d_{20}\}$.
- $CouvPert(g) = \{d_2, d_6\}$.
- $Précision(g) = \frac{2}{10} = 0.2$
- $Rappel(g) = \frac{2}{3} = 0.66$
- $F - mesure(g) = \frac{2 \cdot 0.2 \cdot 0.66}{0.2 + 0.66} = 0.33$

3.6.3.2 Mesures d'évaluation orientées utilisateur

Ces mesures sont principalement issues de réponses à des questionnaires soumis aux collaborateurs afin d'estimer leur engagement dans la session, l'effort cognitif fourni et également le degré de conscience de l'environnement collaboratif. De nombreuses mesures sont définies dans (Shah, 2012), nous présentons seulement les deux principales qui pourraient avoir un impact avec une évaluation d'un modèle de RIC.

L'effort cognitif. La mesure la plus utilisée en RIC est celle liée à la mesure de l'effort cognitif du fait que certains travaux ont démontré que la collaboration engageait un coût additionnel lié aux interactions et à la coordination entre collaborateurs (Fidel et al., 2004). Le formulaire le plus utilisé à cet effet est le formulaire “*TLX instrument*”⁴, défini par la NASA et illustré dans la Figure 3.15. Il inclut six questions qui nécessitent une réponse numérique comprise entre 1 et 5 (ou 7), respectivement très faible à très élevé.

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date
<p>Mental Demand How mentally demanding was the task?</p> <p>Very Low Very High</p>		
<p>Physical Demand How physically demanding was the task?</p> <p>Very Low Very High</p>		
<p>Temporal Demand How hurried or rushed was the pace of the task?</p> <p>Very Low Very High</p>		
<p>Performance How successful were you in accomplishing what you were asked to do?</p> <p>Perfect Failure</p>		
<p>Effort How hard did you have to work to accomplish your level of performance?</p> <p>Very Low Very High</p>		
<p>Frustration How insecure, discouraged, irritated, stressed, and annoyed were you?</p> <p>Very Low Very High</p>		

FIGURE 3.15: Questionnaire du formulaire “TLX instrument”

La communication. Lors d’une tâche de RIC, les interactions entre les utilisateurs peuvent être utilisées pour structurer la collaboration aussi bien que pour échanger des informations. Il est donc important de considérer le processus de communication afin d’évaluer la qualité de la collaboration et les stratégies utilisées. Dans cet objectif, Strijbos et al. (2004) ont défini une catégorisation des messages de communication entre collaborateurs, synthétisée dans la Figure 3.16. Afin de mener une analyse quantitative, González-Ibáñez et al. (2013) proposent de mesurer le nombre de messages générés dans le processus de col-

4. [http : //humansystems.arc.nasa.gov/groups/tlx/downloads/TLXScale.pdf](http://humansystems.arc.nasa.gov/groups/tlx/downloads/TLXScale.pdf)

laboration ainsi que le nombre de mots des messages générés par minute. Ces deux mesures peuvent être calculées sur l'ensemble du groupe ou par utilisateur afin d'identifier les différences entre collaborateurs.

Code	Description	Examples of Statements
Task coordination (TC)	All types of statements regarding coordination, which involve decision making about how the task should be performed	<ol style="list-style-type: none"> 1. We should start writing now 2. Can you search the second one? 3. How do you want to do this? 4. I will work on the reactions and you work on the consequences
Task content (TN)	All types of statements that are related to the content of the task, which include information assessment, layout, structure, and revision of report	<ol style="list-style-type: none"> 1. I found something about consequences 2. Ok, I found stuff on the impact on economy life, people and animals 3. Well, I have how they finally capped off the leak, but I will look up some failed attempts
Task social (TS)	All types of statements that concern group functioning, effort, or attitude as well as opinions in regards to information obtained or information sources	<ol style="list-style-type: none"> 1. This task is really hard 2. We did good 3. Wow, so many animals were killed during the spill 4. I think my answer is the best
Non-task related (NT)	All statements with a social orientation that are not related to the assignment or regarding technical issues of system being used	<ol style="list-style-type: none"> 1. I am hungry 2. How do we save the report? 3. I saw a great movie yesterday 4. What are you going to do tomorrow?
Non-codable (NC)	All types of statements that do not belong any category specified	<ol style="list-style-type: none"> 1. What happened?

FIGURE 3.16: Catégorisation des messages entre collaborateurs (González-Ibáñez et al., 2013)

Résumé

Nous avons présenté dans ce chapitre les concepts de base de la RIC ainsi que les principales approches de l'état de l'art. Ces dernières regroupent des aspects liés au domaine des interactions homme-machine par la proposition d'interfaces adaptées au contexte multi-utilisateurs ainsi que des modèles fondamentaux de RIC dont l'objectif est d'ordonnancer les documents. Enfin, nous avons abordé l'aspect lié à l'évaluation d'un tel domaine en présentant les adaptations des cadres expérimentaux existants en RI individuelle liées à la collaboration.

Discussion et conclusion

Dans cette partie, nous avons proposé une synthèse des travaux de l'état de l'art qui détaille le fondement de nos travaux de recherche. Nous avons abordé, dans un premier temps, les concepts fondamentaux de la RI, communément appelée la RI *ad-hoc* (Salton and McGill, 1983). Puis, nous nous sommes intéressés à la dimension utilisateur d'une session de recherche individuelle en détaillant les approches de la littérature (Agichtein et al., 2006; Balog et al., 2009; Pickens et al., 2008; Amer-Yahia et al., 2007) qui utilisent des facteurs utilisateurs, tels que les intérêts, le niveau d'expertise ou le rôle, dans la fonction de pertinence des documents. Enfin, nous avons axé notre état de l'art vers le domaine de la RIC (Hansen and Järvelin, 2005), sous-jacent aux contributions présentées dans ce manuscrit. Les caractéristiques de chacun de ces sous-domaines de la RI sont détaillées dans le Tableau 3.4. Les principales dimensions d'analyse reposent sur le besoin en information, modélisé par une requête, les utilisateurs, les facteurs utilisés dans l'estimation de la pertinence des documents et les paradigmes sous-jacents au domaine.

	Requête		Utilisateur		Facteurs			Paradigme			
	Individuel	Collaborative	Unique	Groupe	Documents	Individu	Groupe	<i>One size fits all</i>	Contexte	Division du travail	Partage des connaissances
RI ad-hoc	■	□	■	□	■	□	□	■	□	□	□
RI orientée utilisateur	■	■	■	■	■	■	■	□	■	■	■
RI collaborative	■	■	□	■	■	■	■	□	□	■	■

TABLE 3.4: Synthèse des caractéristiques des différents domaines de RI

Plus particulièrement, la RIC peut être abordée selon deux principales approches (Joho et al., 2009) :

- La proposition d’interfaces adaptées au contexte multi-utilisateurs afin de supporter les interactions et la gestion des données liées au cadre de la collaboration.
- La proposition de modèles ou techniques de RI qui permettent d’estimer la pertinence des documents compte tenu du contexte multi-utilisateurs imposé par la collaboration.

Comme présenté précédemment, la RIC est multi-disciplinaire (Foster, 2006; Shah, 2013) recouvrant, entre autres, les aspects liés aux modèles d’ordonnement, à l’analyse du comportement et aux interfaces homme-machine. Ainsi, bien que présentant des démarches intéressantes dans la résolution du problème collaboratif en s’appuyant sur les paradigmes de la division du travail (Kelly and Payne, 2013), du partage des connaissances (Foley and Smeaton, 2010) ou de la conscience de l’environnement collaboratif (Dourish and Bellotti, 1992), les travaux publiés ne couvrent pas l’ensemble du domaine et ouvrent de nombreuses perspectives de recherche. Dans cette optique, nous contribuons dans ce manuscrit à répondre à deux aspects non traités dans les travaux de la littérature :

1. La prise en compte de l’expertise des collaborateurs dans l’objectif d’adapter les modèles d’ordonnement à cette caractéristique. Nous faisons ainsi l’hypothèse que le résultat de la collaboration peut être amélioré si les utilisateurs jugent des documents en accord avec leur niveau d’expertise.
2. La prise en compte de la variabilité des rôles des collaborateurs qui se base sur le fait que les utilisateurs se comportent différemment au cours de la session et peuvent changer de rôle en fonction des documents visités ou des interactions avec les autres membres du groupe. Nous proposons ainsi un nouveau type de modèles de RIC reposant sur une médiation hybride utilisateur-système.

Deuxième partie

Vers des modèles de recherche d'information collaborative basés sur l'expertise

Introduction

D'un point de vue cognitif, l'expertise de domaine est perçue comme la capacité d'un utilisateur à identifier les patrons de compréhension d'une information et à déployer les compétences pour les résoudre (Patel and Arocha J.F., 1999). Plus généralement, l'expertise de domaine d'un utilisateur englobe les connaissances des utilisateurs par rapport à un sujet. Elle est un facteur prépondérant dans le processus de RI dans la mesure où elle infère sur la perception du besoin en information (Hembrooke et al., 2005), la technicité du vocabulaire utilisé (Allen, 1991; White and Dumais, 2009) et la capacité des utilisateurs à identifier des documents pertinents (Hembrooke et al., 2005; White and Dumais, 2009). Par conséquent, il semble important d'adapter les modèles d'ordonnancement par rapport au niveau d'expertise de domaine des utilisateurs.

En RI orientée utilisateur, le facteur d'expertise est principalement utilisé pour la détection d'experts, appelés également autorités du domaine, (Balog et al., 2009; Demartini, 2007; Zhang et al., 2007). Certains travaux (Ben Jabeur et al., 2010; Lee et al., 2001) intègrent l'expertise comme un facteur d'ordonnancement (Bharat and Mihaila, 2001; Kirsch et al., 2006; Ben Jabeur et al., 2010). Cependant, ces travaux sont orientés pour un contexte de RI individuelle, et bien que pouvant être appliqués pour une tâche de RIC, ils ne prennent pas en compte les paradigmes de la collaboration, comme la division du travail. Parmi les approches collaboratives qui proposent une médiation algorithmique orientée collaboration, le facteur d'autorité proposé par Foley and Smeaton (2009), permettant d'agréger les pertinences individuelles afin d'estimer la pertinence collective, qui peut être assimilé à un indicateur du niveau d'expertise des utilisateurs. Cependant, l'intuition de ce modèle basé sur la pertinence collective ne personnalise pas les ordonnancements par rapport à l'expertise des collaborateurs. Une deuxième approche a été utilisée en RIC qui consiste à définir des rôles asymétriques (Golovchinsky et al., 2009b) afin de personnaliser les ordonnancements par rapport aux tâches de recherche assignées aux collaborateurs. Jusqu'alors, les travaux (Pickens et al., 2008; Shah et al., 2010) proposent des paires de rôles, à savoir *prospecteur-mineur* et *rassembleur-surveilleur*, qui sont plutôt basées sur des compétences de recherche, comme la capacité d'un utilisateur à identifier rapidement les documents pertinents ou favoriser la diversité par l'aptitude à reformuler les requêtes.

D'après notre état de l'art du domaine, il n'existe donc pas à ce jour de modèles de RIC permettant à la fois (a) de garantir les paradigmes de la collaboration, et plus particulièrement celui de la division du travail afin d'éviter la redondance entre les actions (Foley and Smeaton, 2010; Kelly and Payne, 2013), et (b) d'intégrer l'expertise de domaine des collaborateurs comme facteur de personnalisation des ordonnancements afin d'optimiser la session de recherche en bénéficiant au mieux de leurs connaissances vis-à-vis du sujet de la requête. La prise en compte des jugements de pertinence dans les modèles de RIC (Foley and Smeaton, 2009; Morris et al., 2008; Pickens et al., 2008; Shah et al., 2010)

est une première approche à cette problématique, mais elle reste superficielle dans la mesure où les jugements de pertinence sont plus abordés comme un outil de recommandation qui restitue aux collaborateurs des documents similaires à ceux précédemment sélectionnés (Foley and Smeaton, 2009). Cette approche s’oppose à l’hypothèse que les modèles d’ordonnancement doivent être capables de former l’utilisateur novice en lui proposant des documents nouveaux et de plus en plus complexes, afin de lui permettre d’augmenter son niveau de connaissances vis-à-vis du sujet (White and Dumais, 2009). Il est cependant possible d’exploiter l’expertise des utilisateurs comme une caractéristique intrinsèque des collaborateurs qui peut être directement impliquée dans le calcul de la pertinence des documents. Par exemple, deux utilisateurs qui collaborent pour résoudre un même besoin en information et ayant des connaissances différentes sur le sujet peuvent apprécier différemment la technicité du vocabulaire ou la spécificité de l’information contenue (Hölscher and Strube, 2000; White and Dumais, 2009). De plus, l’expertise des utilisateurs est d’autant plus pertinente lors d’une session collaborative que la collaboration émerge de la complémentarité des compétences des utilisateurs (Sonnenwald et al., 2004; Spence et al., 2005).

Dans ce contexte, nous proposons de résoudre cette limite en intégrant le facteur d’expertise au sein du processus de RIC, comme un facteur intrinsèque aux utilisateurs dans un modèle d’ordonnancement collaboratif des documents. Plus particulièrement, nous considérons la dimension de l’expertise selon deux axes :

1. L’axe vertical qui fait l’hypothèse que certains collaborateurs ont un niveau d’expertise plus élevé que d’autres vis-à-vis du sujet de la requête, ce qui leur confère respectivement les rôles d’expert et de novice de domaine. Patel and Arocha J.F. (1999) distinguent les experts des novices de par leur mémoire à long et court terme plus élevée pour les informations spécifiques à son domaine de prédilection tandis que les novices ont une vision plus globale du sujet caractérisée seulement par une connaissance superficielle.
2. L’axe horizontal qui fait l’hypothèse que les collaborateurs ont une expertise différente au sein même du sujet et que chacun d’entre eux a des connaissances spécifiques par rapport à des sous-domaines du besoin en information.

Dans le chapitre 4, nous détaillons les deux modèles de RIC sous-jacents aux dimensions verticale et horizontale des niveaux d’expertise des collaborateurs. Nous présentons également le protocole expérimental qui nous a permis de valider ces modèles.

Chapitre 4

Modèles de recherche d'information collaborative basés sur l'expertise des collaborateurs

Introduction

Dans ce chapitre, nous nous intéressons à un contexte de RIC où les utilisateurs sont caractérisés par différents niveaux d'expertise vis-à-vis du besoin en information. Nous proposons deux modèles d'ordonnancement collaboratif de documents intégrant respectivement :

1. Une distinction verticale des niveaux d'expertise sur la base des rôles d'expert et de novice de domaine (Soulier et al., 2014b,c).
2. Une distinction horizontale des niveaux d'expertise sur la base du rôle d'expert de sous-domaines dans l'objectif de répondre à un besoin en information à multiples facettes (Soulier et al., 2013).

La suite du chapitre est organisée comme suit. La section 4.1 présente le contexte et introduit les questions de recherche de nos contributions. La section 4.2 détaille les deux modèles d'ordonnancement collaboratif basés sur l'expertise des utilisateurs. Enfin, l'efficacité de ces modèles est évaluée dans la section 4.3 où nous présentons le cadre expérimental et les résultats obtenus pour chacune des deux contributions.

4.1 Contexte et questions de recherche

La notion d’expertise a été abordée par Shah (2012) dans sa définition de la collaboration de la façon suivante :

“A process involving various agents that may see different aspects of a problem [. . .] (and) can go beyond their own individual expertise.”

Dans ce contexte, la notion d’expertise a été également abordée dans une taxonomie des rôles (Golovchinsky et al., 2009b) qui distingue les niveaux d’expertise entre utilisateurs selon leurs compétences de recherche ou de domaine rassemblant leurs connaissances thématiques du besoin en information. Nous nous intéressons précisément à ce deuxième aspect où les niveaux d’expertise des collaborateurs peuvent être perçus selon deux approches : (a) une distinction verticale des niveaux d’expertise des collaborateurs, leur conférant ainsi les rôles d’expert et de novice de domaine (Soulier et al., 2014b,c) et (b) une distinction horizontale des niveaux d’expertise des collaborateurs les qualifiant ainsi d’expert de sous-domaines (Soulier et al., 2013).

Ainsi, nous proposons d’intégrer le facteur d’expertise de domaine dans l’estimation de la pertinence collaborative des documents en formalisant deux modèles de RIC :

1. Un modèle d’ordonnancement collaboratif basé sur une paire d’utilisateurs caractérisés par des rôles d’expert et de novice de domaine. Ces derniers sont définis sur la base d’une différence de leur niveau d’expertise vis-à-vis du besoin en information partagé. Les niveaux d’expertise peuvent inclure : (a) des différences relatives entre les collaborateurs en considérant qu’un utilisateur est plus expert que l’autre sur un sujet donné ou (b) des différences absolues où le niveau d’expertise est situé aux extrêmes du spectre de l’expertise de domaine permettant d’identifier a priori des rôles d’expert ou de novice. Par convention, nous appelons dans la suite du manuscrit, expert, et respectivement novice, l’utilisateur avec le niveau d’expertise de domaine le plus élevé, respectivement le plus bas. Cette paire de rôles est par exemple illustrée dans quatre domaines d’application, présentés plus en détail dans la section 3.6.3.2 : le domaine médical (McMullan, 2006; ECDPC, 2011), le domaine du e-Discovery (Attfield et al., 2010; Privault et al., 2010), le domaine bibliothécaire (Rudd and Rudd, 1986; Twidale et al., 1997) et académique (Foster, 2006).
2. Un modèle de RIC basé sur un groupe d’experts de sous-domaines dont l’objectif est de résoudre un besoin à multiples facettes. Un exemple courant d’une situation de collaboration entre un groupe d’experts de sous-domaines est celui du domaine médical où des médecins de différentes spécialités peuvent se rencontrer pour décider du processus de soin d’un patient (Samuels, 2011). Les compétences et connaissances des médecins sont complémentaires compte tenu de leur spécialité médicale, ce qui les autorise à aborder la maladie du patient en tenant compte de l’ensemble de ses symptômes. Ainsi, la collaboration permet aux utilisateurs de dépasser l’approche d’analyse surfacique des

résultats qu'un utilisateur unique peut réaliser, en permettant à un ensemble d'utilisateurs de créer une synergie et d'analyser plus en profondeur chacun des aspects du besoin en information.

D'une façon globale, ces deux contributions soulèvent de nombreuses problématiques :

1. *La construction du profil d'expertise des collaborateurs en fonction de la distinction de leur niveau d'expertise.* Plus particulièrement, dans le cas d'une distinction verticale, le profil d'expertise doit permettre (a) de comparer les collaborateurs d'un point de vue de la technicité des termes utilisés au cours de la session (Hölscher and Strube, 2000; White and Dumais, 2009) et (b) d'identifier des documents candidats à l'ordonnancement afin d'augmenter le niveau de connaissance du novice (White and Dumais, 2009). Selon l'approche basée sur la distinction horizontale des niveaux d'expertise, le profil doit permettre d'identifier le degré d'expertise des collaborateurs vis-à-vis des facettes du besoin en information. Ainsi, un simple profil basé sur une distribution multinomiale des termes n'est pas suffisant pour répondre à ce besoin particulier.
2. *L'intégration de l'expertise dans l'estimation de la pertinence des documents par rapport au besoin en information partagé.* Pour les deux approches proposées, la difficulté réside dans le calcul de la pertinence des documents afin de répondre à la fois au besoin en information partagé et aux caractéristiques intrinsèques des collaborateurs afin de personnaliser les ordonnancements tout au long de la session. Dans le cas d'un besoin à multiples facettes, le principal défi repose sur la prise en compte des facettes de la requête dans le modèle d'ordonnancement afin d'apparier le document avec les facettes d'expertise des collaborateurs. Les méthodes de personnalisation en RI individuelle ne sont pas suffisantes dans la mesure où l'insuffisance des compétences et des connaissances d'un utilisateur unique par rapport au besoin en information à multiples facettes peuvent faire obstacle à sa résolution (Shah, 2012) et nous conduisent à proposer une solution basée sur la RIC.
3. *L'ordonnancement collaboratif des documents en fonction de l'expertise des collaborateurs et du besoin partagé.* L'intuition réside dans le fait qu'une simple personnalisation n'est pas suffisante pour garantir l'effet synergique du modèle. Il est nécessaire pour cela de garantir certains paradigmes de la collaboration, et plus particulièrement celui de la division du travail qui est le plus utilisé dans la littérature (Foley and Smeaton, 2009; Kelly and Payne, 2013).

Particulièrement, nos contributions répondent aux questions de recherche suivantes :

1. *En quoi les actions des utilisateurs traduisent-elles leur niveau d'expertise ?* Sur la base des hypothèses proposées par Allen (1991) et White and Dumais (2009), nous considérons que les documents jugés par les utilisateurs reflètent leur niveau d'expertise. Plus particulièrement, afin d'évaluer le niveau d'expertise de la paire de collaborateurs expert/novice, nous proposons

de modéliser leur profil à partir d’une distribution multinomiale des termes pondérés par une mesure relative à la technicité du vocabulaire (Kim, 2006) tandis que nous proposons, en accord avec le besoin à multiples facettes, dans la seconde contribution d’extraire les facettes du besoin en information en utilisant l’algorithme LDA et d’inférer le profil sous la forme d’un vecteur multi-dimensionnel de facettes.

2. *Comment le niveau d’expertise des collaborateurs impacte-t-il l’efficacité de la session ?* D’une façon globale, nous proposons d’injecter le profil des collaborateurs dans le calcul des scores de pertinence des documents. Egalement, dans le cas de la paire expert/novice où la technicité du vocabulaire est prépondérante, nous prenons en compte un facteur de spécificité lié au vocabulaire des utilisateurs. Ainsi, nous proposons un modèle d’ordonnancement collaboratif qui alloue les documents aux utilisateurs par rapport à leur niveau d’expertise ainsi qu’au sujet de la requête collaborative. Pour cela, nous nous assurons de garantir le bénéfice mutuel de la collaboration en allouant les documents aux collaborateurs qui sont les plus à même de les juger en fonction de leur niveau d’expertise de domaine.

4.2 Ordonnancement collaboratif des documents selon des rôles basés sur l’expertise

Dans nos deux contributions, nous considérons un cadre de RIC synchrone impliquant des utilisateurs qui collaborent afin de satisfaire un besoin en information partagé, modélisé par une requête collaborative. Les utilisateurs sont caractérisés par des rôles liés à leur niveau d’expertise vis-à-vis du sujet de la requête. Ces rôles sont symétriques et chaque utilisateur a pour objectif d’identifier les documents pertinents en rapport avec le besoin en information partagé. Le rôle du système de RIC est alors de distribuer les documents aux collaborateurs afin de garantir le bénéfice mutuel qui consiste à satisfaire à la fois le besoin en information collectif, en répondant à la requête collaborative, et les caractéristiques de chacun des membres du groupe, en s’adaptant à leur niveau d’expertise.

Dans ce qui suit, nous présentons le scénario de recherche commun aux deux contributions et détaillons ces deux modèles d’ordonnancement collaboratifs proposés, basés sur une distinction verticale ou horizontale des niveaux d’expertise.

4.2.1 Scénario de recherche

Le scénario de recherche sur lequel repose nos modèles de RIC est illustré dans la Figure 4.1, qui instancie une session sur la base d’une paire d’utilisateurs. Il repose sur une session de recherche S impliquant un ensemble U de m utilisateurs u_j , avec $m \geq 2$. Ces utilisateurs collaborent afin de satisfaire un besoin en

information partagé \mathcal{I} , modélisé par une requête collaborative q . Nous faisons l'hypothèse qu'ils ont des niveaux distincts d'expertise (verticaux ou horizontaux) vis-à-vis du besoin en information et que chaque collaborateur u_j est ainsi caractérisé par un rôle représentatif de son niveau d'expertise, noté $\mathcal{R}(u_j) \in \mathcal{R}$ avec $\mathcal{R} = \{\text{expert de domaine, novice de domaine, expert de sous-domaines}\}$.

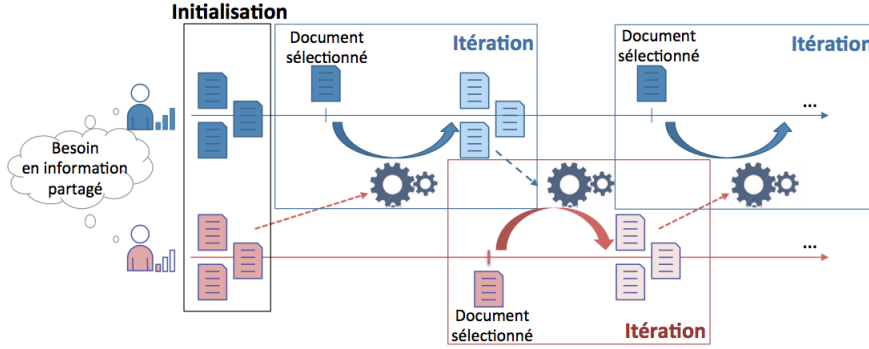


FIGURE 4.1: Exemple d'un scénario de recherche impliquant deux collaborateurs

La session de recherche collaborative S est basée sur un processus itératif de réinjection de pertinence (Foley, 2008). Elle débute par l'expression d'un besoin en information partagé entre les collaborateurs et se termine par l'absence d'actions de jugement (implicites ou explicites) de la part des collaborateurs et/ou l'expression d'un nouveau besoin en information. Entre ces deux événements, la session est caractérisée par deux étapes :

- Une étape d'initialisation où le modèle retourne à chaque utilisateur $u_j \in U$ une liste initiale $l^{init}(u_j)$ de documents distincts et ordonnés en rapport avec le sujet de la requête collaborative q et son rôle $\mathcal{R}(u_j)$.
- Une étape itérative où chaque itération k est lancée par le jugement (ou la sélection) de documents d_i par un utilisateur u_j qui conduit instantanément au réordonnancement des documents non jugés \mathcal{D}_{ns}^k et à la proposition d'une nouvelle liste $l^k(u_j, \mathcal{D}_{ns}^k)$ à ce même utilisateur u_j . Le modèle d'ordonnancement considère le niveau d'expertise de domaine des utilisateurs formulé par leurs jugements de pertinence effectués au cours de la session. L'objectif principal de nos contributions est d'attribuer un document à l'utilisateur qui est le plus à même de le juger en fonction du niveau d'expertise de chaque membre du groupe vis-à-vis du sujet de la requête. De plus, pour respecter le principe de division du travail, seuls les documents non retournés au collaborateur $u_{j'}$ sont contenus dans cette nouvelle liste ordonnée. Réciproquement, le même processus itératif est déployé du côté des autres utilisateurs $u_{j'} \in U$.

4.2.2 Modèle d'ordonnancement collaboratif basé sur les rôles d'expert et de novice de domaine

Dans cette section, nous présentons notre modèle de RIC basé sur la paire de rôles d'expert et de novice de domaine, défini dans la taxonomie de Golovchinsky et al. (2009b). L'objectif de l'expert de domaine est d'affiner ses connaissances par rapport au besoin en information en se concentrant sur des documents spécialisés tandis que le novice a pour objectif d'acquérir une meilleure compréhension du sujet en explorant des documents génériques dont le vocabulaire est limité en terme de technicité.

L'intuition de notre modèle est alors d'ordonner les documents en réponse à un besoin en information partagé en tenant compte des attentes de chacun des rôles ainsi que de leur expertise de domaine. En effet, en accord avec les hypothèses formulées par Allen (1991), Hölscher and Strube (2000) et White and Dumais (2009), les documents dont le vocabulaire est le plus technique doivent être assignés à l'expert alors que le novice doit recevoir des documents moins techniques. De plus, nous souhaitons également répondre à des besoins formulés par White and Dumais (2009) liés à la formation du novice par rapport au sujet du besoin en information.

Dans cet objectif, nous considérons le scénario de recherche illustré dans la Figure 4.1 qui comporte deux événements : l'initialisation et l'itération. L'étape d'initialisation repose sur une simple distribution des résultats les plus pertinents, retournés par la requête, selon un critère de spécificité des documents dans l'objectif de répondre aux caractéristiques des rôles des collaborateurs. Compte tenu des hypothèses sous-jacentes à la paire de rôles liés à la technicité du vocabulaire (Allen, 1991; White and Dumais, 2009), les documents les plus spécifiques parmi les plus pertinents sont retournés à l'expert tandis que les moins spécifiques parmi ce même ensemble sont retournés au novice. Nous proposons de présenter un exemple de cette étape dans l'Illustration 4.1.

Résultant d'un processus plus complexe, nous détaillons une itération de recherche et son modèle d'ordonnancement des documents sous-jacent par la suite. Afin de donner un aperçu des techniques de RI sur lesquelles reposent une itération de recherche, nous proposons de schématiser les étapes de notre modèle dans la Figure 4.2. Nous définissons une itération comme un processus en deux étapes :

1. Une fonction de calcul des scores de pertinence des documents par rapport à chaque rôle qui estiment respectivement la probabilité de pertinence d'un document pour chacun des utilisateurs en fonction de leur rôle d'expert ou de novice de domaine.
2. Une classification des documents vis-à-vis de chacun des deux rôles par un algorithme d'apprentissage afin d'optimiser les scores en utilisant la méthode de maximisation de l'espérance mathématique (EM) (Dempster et al., 1977).

Cette étape permet d'attribuer un document à l'utilisateur qui est le plus à même de le juger compte tenu de son rôle et de son niveau d'expertise.

Illustration 4.1 Initialisation des ordonnancements.

Nous considérons deux utilisateurs u_1 et u_2 qui collaborent pour répondre à une requête collaborative $q = t_1 \wedge t_3$. Leurs rôles sont respectivement $\mathcal{R}(u_1) = \{expert\}$ et $\mathcal{R}(u_2) = \{novice\}$. La collection C comporte 10 documents, soit $C = \{d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8, d_9, d_{10}\}$. L'index I de la collection C inclut 4 termes dont le nombre d'occurrences pour la requête et pour chaque document est exprimé dans le tableau ci-dessous :

	t_1	t_2	t_3	t_4
q	1	0	1	0
d_1	2	3	1	1
d_2	0	0	5	3
d_3	2	1	7	6
d_4	4	1	0	0
d_5	2	0	0	0
d_6	3	0	0	0
d_7	7	1	1	1
d_8	3	3	3	3
d_9	1	4	5	0
d_{10}	0	0	4	0

A partir de ces indications, nous obtenons les vecteurs de poids suivants :

$q = (0.5, 0, 0.5, 0)$;
 $d_1 = (0.29, 0.43, 0.14, 0.14)$
 $d_2 = (0, 0, 0.63, 0.37)$
 $d_3 = (0.12, 0.06, 0.44, 0.28)$
 $d_4 = (0.8, 0.2, 0, 0)$
 $d_5 = (1, 0, 0, 0)$
 $d_6 = (0.3, 0, 0, 0.7)$
 $d_7 = (0.7, 0.1, 0.1, 0.1)$
 $d_8 = (0.25, 0.25, 0.25, 0.25)$
 $d_9 = (0.1, 0.4, 0.5, 0)$
 $d_{10} = (0, 0, 1, 0)$.

Compte tenu du nombre réduit de documents de notre collection C , nous faisons l'hypothèse que seulement 3 documents sont affichés à chaque itération pour chaque utilisateur afin de permettre un renouvellement des listes retournées aux collaborateurs.

Lors de l'étape d'initialisation $k = 0$, nous appliquons le modèle de langue de Jelinek and Mercer (1980) avec un paramètre de lissage d'une valeur arbitraire de $\lambda = 0.8$. Nous obtenons la valeur $RSV(d_i, q_h)$ pour chaque document d_i et faisons l'hypothèse que la spécificité est déjà calculée. Nous la détaillons plus tardivement dans l'illustration 4.2 :

	$RSV(q, d_i)$	$rang(d_i)$	$Spec(d_i)$
d_1	0.24	2	0.19
d_2	0.02	7	0.23
d_3	0.17	3	0.19
d_4	0.03	6	0.15
d_5	0.01	9	0.1
d_6	0.02	8	0.1
d_7	0.10	4	0.19
d_8	0.31	1	0.19
d_9	0.09	5	0.16
d_{10}	0.01	10	0.15

Nous considérons donc les 6 documents les plus pertinents, à savoir $\{d_8, d_1, d_3, d_7, d_9, d_4\}$ et répartissons les trois plus spécifiques à l'expert et, par analogie, les trois documents les moins spécifiques sont attribués au novice. Ainsi, la liste retournée à l'utilisateur u_1 est notée $l^0(u_1, \mathcal{D}_{ns}^0) = \{d_8, d_1, d_3\}$ tandis que celle retournée à l'utilisateur u_2 est notée $l^0(u_2, \mathcal{D}_{ns}^0) = \{d_7, d_9, d_4\}$.

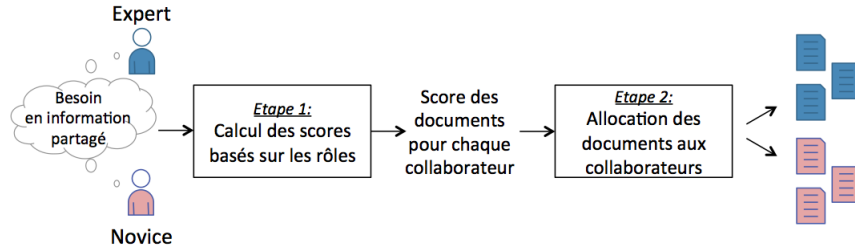


FIGURE 4.2: Aperçu général de notre modèle de RIC basé sur les rôles d’expert et de novice de domaine

Dans ce qui suit, nous présentons les notations relatives à la modélisation des collaborateurs différenciés par des niveaux verticaux de l’expertise de domaine. Puis, nous détaillons les deux étapes principales de notre modèle de RIC associées à une itération de recherche.

4.2.2.1 Notations : modélisation du collaborateur

L’élément central de notre contribution repose sur la modélisation des collaborateurs afin d’être intégrée dans le calcul de la pertinence des documents. D’une façon générale, la pertinence des documents peut être impactée par des caractéristiques reliées à leur rôle, à savoir la technicité du vocabulaire et la nouveauté des documents, dans la session de recherche. Cependant, une autre dimension nous semble intéressante à analyser. En plus d’utiliser des indicateurs numériques liés au niveau d’expertise des utilisateurs, nous souhaitons intégrer dans la fonction de pertinence une dimension qualitative par un profil d’expertise qui recense, sous la forme d’une distribution multinomiale, les termes familiés au collaborateur afin de lui proposer des documents en accord avec son vocabulaire. Nous détaillons ces éléments dans ce qui suit.

Rôles basés sur l’expertise de domaine des collaborateurs. Tout au long de notre contribution, nous nous appuyons sur le fait que des documents spécifiques et nouveaux permettraient aux experts de domaine de découvrir de nouvelles associations sémantiques et d’améliorer leur compréhension structurelle du sujet. A l’inverse, les novices de domaine pourraient acquérir une compréhension globale du sujet afin de découvrir des concepts surfaciques et se construire une connaissance du domaine avec des documents nouveaux et moins spécifiques. Compte tenu de ces hypothèses, nous définissons deux facteurs liés aux rôles d’expert et de novice de domaine :

- Les experts de domaine utilisent un vocabulaire plus technique et une terminologie mieux appropriée durant la session de recherche (Hölscher and Strube, 2000; White and Dumais, 2009). Ainsi, nous proposons de considérer

dans la fonction d'ordonnement le facteur de spécificité des documents $Spec(d_i)$ qui permet d'estimer le niveau de description des termes utilisés dans le document d_i , sur la base de leurs fréquences dans la collection (Kim, 2006). Similairement à la mesure de l'IDF, nous faisons l'hypothèse qu'un terme est spécifique si sa fréquence dans la collection C est faible. Ainsi, le niveau de spécificité $Spec(d_i)$ d'un document d_i est basé sur l'indicateur de spécificité, noté $Pspec$ et détaillé par Kim (2006). Le facteur de spécificité $Spec(d_i)$ d'un document d_i est appliqué sur les termes distincts du document. Comme présenté dans l'illustration 4.2, il est estimé ainsi :

$$Spec(d_i) = avg_{t_v \in d_i} Pspec(t_v) = avg_{t_v \in d_i} \left(\frac{-\log(\frac{fd_{t_v}}{N})}{\alpha} \right) \quad (4.1)$$

où N est la taille de la collection C de documents. fd_{t_v} représente le nombre de documents incluant le terme t_v . α représente la constante entière supérieure à la valeur maximale du numérateur.

Illustration 4.2 Niveau de spécificité d'un document.

Nous considérons la collection C présentée dans l'illustration 4.1.

La spécificité de chaque document est calculée comme suit :

- Si un terme n'apparaît qu'une fois dans la collection, alors $-\log(1/10) = 2.30$; nous posons alors $\alpha = 3$.
- $d_1 = \frac{\frac{-\log(\frac{8}{10})}{3} \frac{-\log(\frac{6}{10})}{3} \frac{-\log(\frac{7}{10})}{3} \frac{-\log(\frac{5}{10})}{3}}{4} = 0.19$
- $d_2 = \frac{\frac{-\log(\frac{7}{10})}{3} \frac{-\log(\frac{5}{10})}{3}}{2} = 0.23$
- $d_3 = \frac{\frac{-\log(\frac{8}{10})}{3} \frac{-\log(\frac{6}{10})}{3} \frac{-\log(\frac{7}{10})}{3} \frac{-\log(\frac{5}{10})}{3}}{4} = 0.19$
- $d_4 = \frac{\frac{-\log(\frac{8}{10})}{3} \frac{-\log(\frac{6}{10})}{3}}{2} = 0.15$
- $d_5 = \frac{\frac{-\log(\frac{8}{10})}{3}}{1} = 0.01$
- $d_6 = \frac{\frac{-\log(\frac{8}{10})}{3}}{1} = 0.1$
- $d_7 = \frac{\frac{-\log(\frac{8}{10})}{3} \frac{-\log(\frac{6}{10})}{3} \frac{-\log(\frac{7}{10})}{3} \frac{-\log(\frac{5}{10})}{3}}{4} = 0.19$
- $d_8 = \frac{\frac{-\log(\frac{8}{10})}{3} \frac{-\log(\frac{6}{10})}{3} \frac{-\log(\frac{7}{10})}{3} \frac{-\log(\frac{5}{10})}{3}}{4} = 0.19$
- $d_9 = \frac{\frac{-\log(\frac{8}{10})}{3} \frac{-\log(\frac{6}{10})}{3} \frac{-\log(\frac{7}{10})}{3}}{3} = 0.16$
- $d_{10} = \frac{\frac{-\log(\frac{7}{10})}{3}}{1} = 0.15$

- La nouveauté des documents est un facteur important dans l'évaluation d'un ordonnancement (Soboroff and Harman, 2005), d'autant plus dans un processus de recherche par réinjection de pertinence afin d'éviter de retourner des documents similaires à ceux déjà jugés. De plus, la nouveauté offre aux

collaborateurs la possibilité d'acquérir de nouvelles connaissances (White and Dumais, 2009). Dans ce contexte, nous définissons le facteur de nouveauté des documents $Nouv(d_i, \mathcal{D}(u_j)^k)$ qui estime le gain marginal de connaissances qu'un document d_i peut apporter à l'utilisateur u_j compte tenu des documents $\mathcal{D}(u_j)^k$ qu'il a déjà sélectionnés au cours de la session à l'itération k . Nous utilisons la mesure de nouveauté basée sur la distance entre les items – "Distance-based Item Novelty" – (Castells et al., 2011) appliquée à l'ensemble de documents sélectionnés $\mathcal{D}(u_j)^k$ par l'utilisateur u_j . Le niveau de nouveauté $Nouv(d_i, \mathcal{D}(u_j)^k)$ du document d_i compte tenu de l'ensemble des documents précédemment sélectionnés $\mathcal{D}(u_j)^k$ par l'utilisateur u_j à l'itération k est estimé de la façon suivante (comme présenté dans l'Illustration 4.3) :

$$Nouv(d_i, \mathcal{D}(u_j)^k) = \min_{d_{i'} \in \mathcal{D}(u_j)^k} d(d_i, d_{i'}) \quad (4.2)$$

où la distance $d(d_i, d_{i'})$ entre deux documents d_i et $d_{i'}$ dépend de la fonction de similarité de Jaccard $sim(d_i, d_{i'}) = \frac{|d_i \cap d_{i'}|}{|d_i \cup d_{i'}|}$ avec $|d_i \cap d_{i'}|$, respectivement $|d_i \cup d_{i'}|$, qui représentent le nombre de termes inclus dans les documents d_i et $d_{i'}$, respectivement d_i ou $d_{i'}$. Au final, la distance d est estimée par $d(d_i, d_{i'}) = 1 - sim(d_i, d_{i'})$.

Illustration 4.3 Niveau de nouveauté d'un document.

Nous considérons la collection C présentée dans l'Illustration 4.1. L'utilisateur u_1 a sélectionné trois documents $\{d_1, d_2, d_3\}$ entre le début de la session et l'itération $k = 5$. L'ensemble des documents sélectionnés par l'utilisateur u_1 est donc noté $\mathcal{D}(u_1)^5 = \{d_1, d_2, d_3\}$. La nouveauté $Nouv(d_4, \mathcal{D}(u_1)^5)$ du document d_4 compte tenu de cet ensemble $\mathcal{D}(u_1)^5$ est estimée comme suit :

$$\min(d(d_1, d_4); d(d_2, d_4); d(d_3, d_4)) = \min(1 - \frac{6}{12}; 1 - \frac{0}{13}; 1 - \frac{6}{21}) = 0.5$$

Profil d'expertise des collaborateurs. Le profil d'un collaborateur est associé à une approche qualitative du niveau de connaissances des collaborateurs dans la mesure où il recense le vocabulaire pour lequel l'utilisateur est le plus familier. Il peut être utilisé pour ordonner collaborativement les documents tout en personnalisant les listes retournées à l'utilisateur. Ainsi, un profil d'expertise $\pi(u_j)^k$ est associé à chaque utilisateur u_j selon son expertise de domaine vis-à-vis du besoin en information partagé, en utilisant une modélisation basée sur la fréquence des termes inclus dans les documents jugés par l'utilisateur à l'itération k . Compte tenu du scénario de recherche de notre modèle, nous considérons une représentation dynamique du profil d'expertise $\pi(u_j)^k$ d'un utilisateur u_j qui est automatiquement mis à jour à chaque itération k . L'expertise d'un utilisateur u_j est extraite de la requête collaborative q et des documents $\mathcal{D}(u_j)^k$ sélectionnés par l'utilisateur u_j depuis le début de la session S jusqu'à l'itération de recherche k . Le poids w_{vj}^k d'un terme t_v pour le profil d'expertise $\pi(u_j)^k$ correspond à la moyenne des poids de ce terme dans la requête q ainsi que dans chaque docu-

ment sélectionné $d_i \in \mathcal{D}(u_j)^k$. Le profil d'un collaborateur est modélisé de la façon suivante :

$$\begin{aligned} \pi(u_j)^k &= \{(t_1, w_{1j}^k), \dots, (t_v, w_{vj}^k), \dots, (t_V, w_{Vj}^k)\} \\ \text{avec } w_{vj}^k &= \frac{w_{vq} + \sum_{d_i \in \mathcal{D}(u_j)^k} w_{vi}}{1 + |\mathcal{D}(u_j)^k|} \end{aligned} \quad (4.3)$$

où les poids w_{vq} et w_{vi} sont respectivement ceux du terme t_v dans la requête q et le document d_i . Le nombre de documents sélectionnés par l'utilisateur u_j à l'itération k est noté $|\mathcal{D}(u_j)^k|$. Nous modélisons également la requête q de façon vectorielle $q = \{(t_1, w_{1q}), \dots, (t_v, w_{vq}), \dots, (t_V, w_{Vq})\}$ où chaque couple (t_v, w_{vq}) correspond à un terme t_v de la requête et son poids associé w_{vq} . Le profil est normalisé de façon à obtenir $\sum_{v \in \mathcal{I}} w_{vj}^k = 1$. Notons qu'en début de session, le profil est initialisé seulement à partir des poids des termes de la requête q . Un exemple de modélisation du profil est présenté dans l'Illustration 4.4.

Illustration 4.4 Construction du profil d'expertise d'un collaborateur.

Nous considérons les notations présentées dans l'Illustration 4.1.

Nous souhaitons construire le profil de l'utilisateur u_2 sachant qu'il a sélectionné les documents d_4 et d_5 , respectivement aux itérations 1 et 2.

Itération 0 : Le profil $\pi(u_2)^0$ est initialisé par rapport à la requête collaborative, soit $\pi(u_2)^0 = (0.5, 0, 0.5, 0)$.

Itération 1 : Le profil est mis à jour par rapport au document d_4 :
 $\pi(u_2)^1 = (\frac{0.5+0.8}{2}, \frac{0.2}{2}, \frac{0.5}{2}, 0) = (0.65, 0.1, 0.25, 0)$.

Itération 2 : Le profil est mis à jour par rapport au document d_5 :
 $\pi(u_2)^2 = (\frac{0.5+0.8+1}{3}, \frac{0.2}{3}, \frac{0.25}{3}, 0) = (0.77, 0.07, 0.16, 0)$.

4.2.2.2 Calcul des scores de pertinence des documents basés sur les rôles

L'objectif de cette étape est d'estimer la pertinence des documents pour chacun des collaborateurs en fonction du sujet de la requête collaborative et de leurs caractéristiques liées à leur niveau d'expertise vis-à-vis de ce même sujet. Ainsi, la pertinence $P^k(d_i|u_j, q)$ d'un document d_i vis-à-vis de la requête collaborative q , par rapport à un utilisateur u_j pour l'itération de recherche k , est estimée comme présenté dans l'Equation 4.4 :

$$P^k(d_i|u_j, q) = \frac{P^k(u_j|d_i, q) \cdot P^k(d_i|q)}{P^k(u_j|q)} \quad (4.4)$$

En se basant sur les hypothèses que la probabilité $P^k(u_j|q)$ n'est pas discriminante dans l'estimation de la pertinence du document et que l'utilisateur u_j est

indépendant de la requête q , on obtient :

$$P^k(d_i|u_j, q) \propto P^k(u_j|d_i) \cdot P^k(d_i|q) \quad (4.5)$$

D'une part, nous proposons tout d'abord de calculer la probabilité $P^k(d_i|q)$ en utilisant les probabilités conditionnelles, comme présenté dans l'Equation 4.6. Ensuite, en considérant que le score du document par rapport à la requête est invariant quelle que soit l'itération k et que la probabilité $P(d_i)$ n'est pas discriminante dans le calcul de la probabilité $P^k(d_i|q)$, la probabilité peut donc être réécrite de la façon suivante (Equation 4.7) :

$$P^k(d_i|q) = P(q|d_i) \cdot P(d_i) \quad (4.6)$$

$$\propto P(q|d_i) \quad (4.7)$$

Dans l'objectif d'estimer la probabilité $P(q|d_i)$, nous faisons l'hypothèse que le document est représenté par une distribution multinomiale des termes θ_{d_i} . Par conséquence, la probabilité $P(q|d_i)$ correspond à la probabilité $P(q|\theta_{d_i})$ estimée par le biais du modèle de langue lissé (Jelinek and Mercer, 1980) :

$$P(q|\theta_{d_i}) = \prod_{(t_v, w_{vq}) \in q} [\lambda P(t_v|\theta_{d_i}) + (1 - \lambda)P(t_v|\theta_C)]^{w_{vq}} \quad (4.8)$$

$$\text{avec } P(t_v|\theta_{d_i}) = \frac{tf(t_v, d_i)}{|d_i|} \text{ et } P(t_v|\theta_C) = \frac{tf(t_v, C)}{|C|}$$

où le document d_i est représenté par une distribution multinomiale θ_{d_i} par rapport aux termes, $tf(t_v, d_i)$ et $tf(t_v, C)$ sont respectivement les fréquences des termes t_v dans le document d_i et la collection de document C . w_{vq} représente le poids du terme t_v dans la requête q . Le nombre de termes dans le document d_i et la collection C sont notés $|d_i|$ et $|C|$, λ est un paramètre de lissage.

D'autre part, la probabilité $P^k(u_j|d_i)$ est estimée suivant une modélisation basée sur le modèle de langue par la probabilité $P(\pi(u_j)^k|\theta_{d_i})$ où l'utilisateur u_j est estimé par sa distribution des termes de son profil $\pi(u_j^k)$ à l'itération k selon le modèle de langue θ_{d_i} du document d_i . Elle est également estimée en utilisant l'approche de Jelinek and Mercer (1980) :

$$P(\pi(u_j)^k|\theta_{d_i}) = \prod_{(t_v, w_{vj}^k) \in \pi(u_j)^k} [\lambda_{ij}^k P(t_v|\theta_{d_i}) + (1 - \lambda_{ij}^k)P(t_v|\theta_C)]^{w_{vj}^k} \quad (4.9)$$

où λ_{ij}^k est un paramètre de lissage dynamique basé sur le document d_i et le rôle $\mathcal{R}(u_j)$ de l'utilisateur u_j d'expert ou de novice. Pour l'expert u_j et étant donné un document d_i , la valeur de λ_{ij} est importante si le document est spécifique et nouveau par rapport aux documents déjà sélectionnés. Pour le novice, cette valeur est importante si le document d_i est peu spécifique et nouveau par rapport aux documents déjà sélectionnés. Par conséquent, nous estimons λ_{ij} comme suit, en considérant $\beta = 1$ pour l'expert et $\beta = -1$ pour le novice de domaine :

$$\lambda_{ij}^k = \frac{Now(d_i, \mathcal{D}(u_j)^k) \cdot Spec(d_i)^\beta}{\max_{d_{i'} \in \mathcal{D}} Now(d_i, \mathcal{D}(u_j)^k) \cdot Spec(d_{i'})^\beta} \quad (4.10)$$

Illustration 4.5 Scores des documents selon les rôles des collaborateurs.

Nous considérons les notations présentées dans l'Illustration 4.1 ainsi que les calculs effectués dans les exemples précédents. Soit l'itération $k = 1$ où l'utilisateur u_2 sélectionne le document d_4 . Le calcul du score est seulement déroulé pour le document d_1 .

Collaborateur u_1 dont le profil est $\pi(u_1)^1 = (0.5, 0, 0.5, 0)$ et l'ensemble des documents sélectionnés est $\mathcal{D}(u_1)^1 = \{\}$:

1. Le score $P^1(d_1|q) = 0.24$ est estimé dans l'Exemple 4.1. Ce score est invariant des itérations.
2. Le score $P^1(u_1|d_1)$ dépend du paramètre de lissage $\lambda_{11}^1 = \frac{1*0.19}{0.23} = 0.85$ où 0.19 représente la spécificité du document d_1 , 1 est le score de nouveauté du document par rapport à l'ensemble vide \mathcal{D}_{ns}^1 et 0.23 représente la valeur maximale du produit $Nouv(d_1, \mathcal{D}(u_1)^1) * Spec(d_1)$ sur l'ensemble des documents $d_i \in \mathcal{D}_{ns}^1$.
Ainsi, $P^1(u_1|d_1) = (0.85 * \frac{2}{7} + 0.15 * \frac{24}{84})^{0.05} + (0.85 * \frac{3}{7} + 0.15 * \frac{13}{84})^0 + (0.85 * \frac{1}{7} + 0.15 * \frac{26}{84})^{0.05} + (0.85 * \frac{1}{7} + 0.15 * \frac{21}{84})^0 = 0.22$
3. Le score final est $P^1(d_1|u_1) = 0.24 * 0.22 = 0.05$.

Collaborateur u_2 dont le profil est $\pi(u_2)^1 = (0.65, 0.1, 0.25, 0)$ et dont l'ensemble des documents sélectionnés est $\mathcal{D}(u_2)^1 = \{d_4\}$:

1. Le score $P^1(d_1|q) = 0.24$ est estimé dans l'Exemple 4.1. Ce score est invariant des itérations.
2. Le score $P^1(u_2|d_1)$ dépend du paramètre de lissage $\lambda_{12}^1 = \frac{0.5 * \frac{1}{0.19}}{4.62} = 0.56$ où 0.5 est le score de nouveauté du document par rapport à \mathcal{D}_{ns}^1 et 4.62 représente la valeur maximale du produit $Nouv(d_1, \mathcal{D}(u_2)^1) * Spec(d_1)^{-1}$ sur l'ensemble des documents $d_i \in \mathcal{D}_{ns}^1$.
Ainsi, $P^1(u_2|d_1) = (0.56 * \frac{2}{7} + 0.44 * \frac{24}{84})^{0.65} + (0.56 * \frac{3}{7} + 0.44 * \frac{13}{84})^{0.1} + (0.56 * \frac{1}{7} + 0.44 * \frac{26}{84})^{0.25} + (0.56 * \frac{1}{7} + 0.44 * \frac{21}{84})^0 = 0.22$
3. Le score final est $P^1(d_1|u_1) = 0.24 * 0.22 = 0.05$.

L'ensemble des scores normalisés des documents non sélectionnés $\mathcal{D}_{ns}^1 = \{d_1, d_2, d_3, d_5, d_6, d_7, d_8, d_9, d_{10}\}$ vis-à-vis de chacun des collaborateurs est présenté dans le tableau suivant :

	$P^1(d_i u_1)$	$P^2(d_i u_1)$
d_1	0.23	0.28
d_2	0	0.03
d_3	0.16	0.11
d_5	0.01	0.01
d_6	0.03	0.02
d_7	0.12	0.14
d_8	0.34	0.34
d_9	0.10	0.06
d_{10}	0.01	0.01

4.2.2.3 Apprentissage des ordonnancements de documents basé sur les rôles d'expert et de novice de domaine

Lors de l'étape précédente, nous avons estimé le score personnalisé des documents vis-à-vis de chacun des collaborateurs et de la requête. Dans le but d'optimiser ces scores, nous utilisons l'algorithme de maximisation de l'espérance (EM) (Dempster et al., 1977; Marsland, 2009) en tenant compte de deux distributions de probabilité reconnues (Manmatha et al., 2001), à savoir les lois gaussiennes et exponentielles qui modélisent respectivement les fonctions de pertinence et de non-pertinence des documents. Le modèle complet d'optimisation repose en conséquence sur deux probabilités $P(R_{ij} = 1|x_{ij}^k)$ et $P(R_{ij} = 0|x_{ij}^k)$ qui estiment respectivement la pertinence et la non pertinence du document d_i par rapport à son score x_{ij}^k pour l'iteration k . Ce score traduit la probabilité de pertinence d'un document d_i en fonction de la requête q et du rôle $\mathcal{R}(u_j)$ de l'utilisateur u_j , dont l'estimation est présentée dans l'équation 4.4 (cf. Section 4.2.2.2). Dans cet objectif, l'algorithme de maximisation de l'espérance (EM) est un algorithme itératif et se décompose en deux étapes. La première, appelée *Etape-E*, estime la probabilité $P(R_{ij} = 1|x_{ij}^k)$ selon un modèle de mélange qui considère, d'une part, une loi de probabilité gaussienne et, d'autre part, une loi de probabilité exponentielle. Nous utilisons seulement la probabilité $P(R_{ij} = 1|x_{ij}^k)$ compte tenu de notre objectif d'identifier le collaborateur le plus à même de juger de la pertinence d'un document. La seconde étape, appelée *Etape-M*, met à jour les paramètres du modèle de mélange en maximisant la vraisemblance de la probabilité. Les résultats de cet algorithme sont ensuite utilisés pour ordonnancer et classer les documents par rapport à l'utilisateur le plus à même de les juger. Les notations utilisées sont décrites dans le Tableau 4.2 et le principe d'ordonnement détaillé dans l'Algorithme 1.

Ci-dessous, nous détaillons les deux principales étapes de l'algorithme utilisé pour l'allocation des documents par méthode d'apprentissage. Nous présentons un exemple dans l'Illustration 4.6.

1. **Apprentissage de l'appariement document-collaborateur.** Nous appliquons l'algorithme de maximisation de l'espérance EM dont les deux étapes sont les suivantes :

- *Etape-E.* Lors de cette étape, nous estimons la probabilité $P(R_{ij} = 1|x_{ij}^k)$ de la pertinence d'un document pour un utilisateur u_j sachant son score x_{ij}^k en appliquant le théorème de Bayes comme suit :

$$\begin{aligned} P(R_{ij} = 1|x_{ij}^k) &= \frac{P(x_{ij}^k|R_{ij}=1) \cdot P(R_{ij}=1)}{P(x_{ij}^k)} \\ &= \frac{P(x_{ij}^k|R_{ij}=1) \cdot P(R_{ij}=1)}{P(x_{ij}^k|R_{ij}=1) \cdot P(R_{ij}=1) + P(x_{ij}^k|R_{ij}=0) \cdot P(R_{ij}=0)} \end{aligned} \quad (4.11)$$

La probabilité $P(R_{ij} = 1|x_{ij}^k)$ peut être réécrite en utilisant le modèle de mélange basé sur les lois gaussiennes et exponentielles :

$$P(R_{ij} = 1|x_{ij}^k) \propto \frac{\alpha_j^k \cdot \phi_j^k(x_{ij}^k)}{\alpha_j^k \cdot \phi_j^k(x_{ij}^k) + (1 - \alpha_j^k) \cdot \psi_j^k(x_{ij}^k)} \quad (4.12)$$

$X_j^k = \{x_{1j}^k, \dots, x_{ij}^k, \dots, x_{nsj}^k\}$	Le score des documents non sélectionnés \mathcal{D}_{ns} où chaque élément x_{ij} est estimé par l'équation 4.4.
X^k	La matrice des scores x_{ij}^k pour l'ensemble des documents d_i et utilisateurs u_j .
$R_{ij} = \{0, 1\}$	La variable latente de la non pertinence ou la pertinence d'un document d_i pour un utilisateur u_j .
ϕ_j^k	La fonction de densité de la loi gaussienne modélisant la pertinence des documents par rapport à l'utilisateur u_j à l'itération k .
ψ_j^k	La fonction de densité de la loi exponentielle modélisant la non pertinence des documents par rapport à l'utilisateur u_j à l'itération k .
α_j^k	Le coefficient du modèle de mélange.
θ_j^k	Les paramètres des distributions des scores à l'itération k étant donné l'utilisateur u_j correspondant à la loi gaussienne ϕ_j^k , μ_j^k et σ_j^k , et à la loi exponentielle ψ_j^k , λ_j^k .

TABLE 4.1: Notations utilisées pour l'algorithme d'allocation à la paire d'expert et de novice de domaine par méthode d'apprentissage

- *Etape-M*. Cette étape permet de mettre à jour les paramètres θ_j^k et définit la convergence de l'algorithme. Pour cela, nous utilisons la log-vraisemblance des données complétées $\mathcal{L}(R_{ij} = 1 | X_j^k, \theta_j^k)$, estimée de la façon suivante :

$$\mathcal{L}(R_{ij} = 1 | X_j^k, \theta_j^k) = \sum_{h=1}^n \sum_{j=1}^2 \log(P(x_{ij}^k, R_{ij} = 1 | \theta_j^k)) P(R_{ij} = 1 | x_{ij}^k)$$

$$\text{avec} \quad P(x_{ij}^k, R_{ij} = 1 | \theta_j^k) = \alpha_j^k \cdot \phi_j^k(x_{ij}^k) \quad (4.13)$$

Nous obtenons pour chaque utilisateur u_j une liste δ_j^k de documents ordonnés selon la probabilité $P(R_{ij} = 1 | x_{ij}^k)$.

- 2. Allocation des documents aux collaborateurs.** L'objectif, ici, est de déterminer quel utilisateur est le plus à même de juger de la pertinence d'un document d_i . Nous utilisons les listes δ_j^k et $\delta_{j'}^k$, de chaque utilisateur et, plus particulièrement, nous observons les rangs du document d_i dans ces deux listes. Intuitivement, nous faisons l'hypothèse que si le rang d'un document est plus petit dans la liste δ_j^k , alors ce document est présenté à l'utilisateur u_j . Par conséquent, nous modélisons la fonction d'allocation suivante :

$$r_{jj'}^k(d_i, \delta_j^k, \delta_{j'}^k) = \begin{cases} 1 & \text{si } rang(d_i, \delta_j^k) < rang(d_i, \delta_{j'}^k) \\ 0 & \text{sinon} \end{cases} \quad (4.14)$$

où $rang(d_i, \delta_j^k)$ et $rang(d_i, \delta_{j'}^k)$ représentent le rang du document d_i respectivement dans les listes ordonnées δ_j^k et $\delta_{j'}^k$. Ainsi, si $r_{jj'}^k(d_i, \delta_j^k, \delta_{j'}^k) = 1$,

Algorithm 1: Allocation des documents par méthode d'apprentissage

Données: $\mathcal{D}_{ns}^k, U, u_j, X^k$ **Résultat:** $l^k(u_j, \mathcal{D}_{ns}^k)$ **Début**

/* Etape 1: Apprentissage de l'appariement document-collaborateur */

Répéter $\delta_j^k = \emptyset$ $\delta_{j'}^k = \emptyset$

/* Etape-E */

Pour chaque document $d_i \in \mathcal{D}_{ns}$ **faire****Pour chaque utilisateur** $u_{j'}$ **faire**

$$P(R_{ij'} = 1 | x_{ij'}^k) = \frac{\alpha_{j'}^k \cdot \phi_{j'}^k(x_{ij'}^k)}{\alpha_{j'}^k \cdot \phi_{j'}^k(x_{ij'}^k) + (1 - \alpha_{j'}^k) \cdot \psi_{j'}^k(x_{ij'}^k)}$$
$$\delta_{j'}^k = \delta_{j'}^k \cup (d_i, P(R_{ij'} = 1 | x_{ij'}^k))$$

/* Etape-M */

Pour chaque utilisateur $u_{j'}$ **faire**

$$S_{j'}^k = \sum_{d_{i'} \in \mathcal{D}_{ns}^k} P(R_{ij'} = 1 | x_{i'j'}^k)$$

$$\alpha_{j'}^k = \frac{1}{n} S_{j'}^k$$

$$\mu_{j'}^k = \frac{1}{S_{j'}^k} \sum_{d_{i'} \in \mathcal{D}_{ns}^k} P(R_{ij'} = 1 | x_{i'j'}^k) \cdot x_{i'j'}^k$$

$$\sigma_{j'}^k = \frac{1}{S_{j'}^k} \sum_{d_{i'} \in \mathcal{D}_{ns}^k} P(R_{ij'} = 1 | x_{i'j'}^k) \cdot (x_{i'j'}^k - \mu_{j'}^k)^2$$

$$\lambda_{j'}^k = \frac{\sum_{d_{i'} \in \mathcal{D}_{ns}^k} (1 - P(R_{ij'} = 1 | x_{i'j'}^k))}{\sum_{d_{i'} \in \mathcal{D}_{ns}^k} (1 - P(R_{ij'} = 1 | x_{i'j'}^k)) \cdot x_{i'j'}^k}$$

jusqu'à convergence

/* Etape 2: Allocation des documents aux utilisateurs */

Pour chaque document $d_i \in \mathcal{D}_{ns}$ **faire****Si** $r_{jj'}(d_i, \delta_j^k, \delta_{j'}^k)^k = 1$ **alors**

$$l^k(u_j, \mathcal{D}_{ns}^k) = l^k(u_j, \mathcal{D}_{ns}^k) \cup (d_i, P(R_{ij} = 1 | x_{ij}^k))$$

Retourne $l^k(u_j, \mathcal{D}_{ns}^k)$

le document d_i est associé à l'utilisateur u_j pour construire sa liste finale $l^k(u_j, \mathcal{D}_{ns}^k)$; sinon, il est présenté à l'utilisateur $u_{j'}$.

Enfin, pour garantir la division du travail, un document d_i est inclus dans la liste finale $l^k(u_j, \mathcal{D}_{ns}^k)$ de l'utilisateur u_j si ce document n'est pas présenté dans la liste $l^k(u_{j'}, \mathcal{D}_{ns}^k)$ actuellement présentée à l'itération k à l'autre utilisateur $u_{j'}$. Notons que l'algorithme EM participe également à l'application de la division du travail dans la mesure où les documents sont classifiés selon les collaborateurs, afin d'attribuer un document à un unique collaborateur qui est le plus à même de le juger.

Illustration 4.6 Ordonnancement collaboratif des documents.

Nous considérons les notations présentées dans l'illustration 4.1 et les résultats obtenus dans l'illustration 4.5. A titre d'illustration, nous effectuons seulement le calcul pour deux itérations de l'algorithme EM.

Pour l'optimisation du score du document d_1 vis-à-vis de l'utilisateur u_1 à l'itération 1, par exemple, $x_{11}^1 = \frac{0.5 \cdot \mathcal{N}(0.23, 0.11, 0.12)}{0.5 \cdot \mathcal{N}(0.23, 0.11, 0.12) + 0.5 \cdot \mathcal{E}(0.23, 9)} = 0.64$.

Pour l'itération 2, $x_{11}^2 = \frac{0.5 \cdot \mathcal{N}(0.64, 1.35, 13.09)}{0.5 \cdot \mathcal{N}(0.64, 1.35, 13.09) + 0.5 \cdot \mathcal{E}(0.64, 11.90)} = 0.004$.

Les résultats finaux sont synthétisés dans le tableau suivant :

		Itération k=0		Itération k=1		Itération k=2	
		x_{i1}^0	x_{i2}^0	x_{i1}^1	x_{i2}^1	x_{i1}^2	x_{i2}^2
Etape-E	d_1	0.23	0.27	0.64	0.64	0.004	0.005
	d_2	0	0.03	0.19	0.28	0.69	0.26
	d_3	0.16	0.11	0.59	0.50	0.01	0.02
	d_5	0.01	0.01	0.22	0.22	0.52	0.61
	d_6	0.03	0.02	0.27	0.26	0.28	0.33
	d_7	0.13	0.14	0.54	0.55	0.01	0.01
	d_8	0.34	0.34	0.54	0.55	0.01	0.01
	d_9	0.10	0.06	0.48	0.37	0.02	0.09
	d_{10}	0.01	0.01	0.21	0.21	0.56	0.58
	μ_j^k	0.11	0.11	1.35	1.36	2.02	2.18
Etape-M	σ_j^k	0.12	0.12	13.09	13.35	29.22	30.00
	λ_j^k	9	9	11.90	11.86	2.14	2.27
	α_j^k	0.5	0.5	0.41	0.40	0.23	0.21
	vraiss.			-5.451		-4.03	

Nous supposons que les résultats obtenus sont optimisés et que l'algorithme EM a atteint la convergence à l'itération 2. Nous pouvons attribuer un rang aux documents, comme indiqué dans le tableau ci-dessous :

	$rang(d_i, \delta_1^1)$	$rang(d_i, \delta_2^1)$	$r_{21}^1(d_i, \delta_2^1, \delta_1^1)$
d_1	9	9	0
d_2	1	4	0
d_3	8	6	1
d_5	3	1	1
d_6	4	3	1
d_7	6	7	0
d_8	7	8	0
d_9	5	5	0
d_{10}	2	2	0

Par la comparaison des rangs, on note que seuls les documents $\{d_3, d_5, d_6\}$ sont candidats pour être retournés à l'utilisateur u_2 . Cependant, dans l'objectif de garantir la division du travail, le document d_3 ne peut être présenté à l'utilisateur car il est couramment présenté à l'utilisateur u_1 . Au final, la liste $l^1(u_2, \mathcal{D}_{ns}^1) = \{d_5, d_6\}$.

4.2.3 Modèle d'ordonnancement collaboratif basé sur les rôles d'experts de sous-domaines

Dans cette section, nous présentons une extension du modèle de RIC basé sur les rôles d'expert et de novice de domaine afin de considérer les niveaux d'expertise entre collaborateurs selon une distinction horizontale. Plus particulièrement, nous faisons l'hypothèse que des utilisateurs avec des compétences de sous-domaines peuvent collaborer afin de résoudre un besoin à multiples facettes. Le principal objectif de notre contribution est d'ordonner dynamiquement et itérativement les documents en tenant compte d'un besoin à multiples facettes, des expertises de sous-domaines des utilisateurs ainsi que du paradigme de la division du travail. Similairement à White and Dumais (2009), nous faisons l'hypothèse que les niveaux d'expertise, même considérés selon une distinction horizontale, impactent sur le comportement des utilisateurs et l'efficacité de la session. Par conséquent, il nous semble nécessaire de considérer cet aspect dans la fonction d'ordonnancement collaboratif des documents.

Notre modèle repose sur le scénario de recherche décrit dans la Section 4.2.1. Etant donné la considération d'un besoin en information à multiples facettes, nous procédons par une étape d'initialisation, détaillée dans la section 4.2.3.2, qui consiste en l'apprentissage des facettes de la requête au travers de la diversification des documents (Carbonell and Goldstein, 1998) et un algorithme d'extraction thématique, à savoir l'algorithme LDA (Blei et al., 2003).

Ensuite, comme indiqué dans la Figure 4.3, nous proposons pour chaque itération du processus de réinjection de pertinence une approche d'ordonnancement en deux étapes :

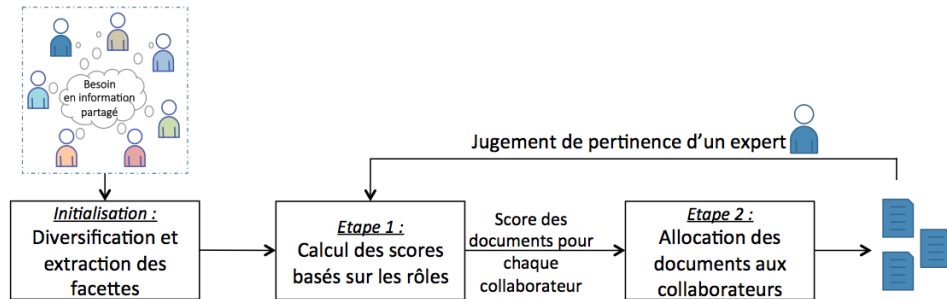


FIGURE 4.3: Aperçu général de notre modèle de RIC basé sur les rôles d'experts de sous-domaines

- *Etape 1* : Un calcul des scores qui estime la pertinence des documents par rapport à chacun des experts de sous-domaines en tenant compte de leur profil d'expertise et du sujet de la requête collaborative (section 4.2.3.3).

- *Etape 2* : Une classification des documents vis-à-vis de chacun des collaborateurs afin d’optimiser les scores des documents en utilisant la méthode de maximisation de l’espérance (Dempster et al., 1977). Cette étape a pour objectif d’assigner un document à l’utilisateur qui est le plus à même de juger du contenu du document selon son (ou ses) sous-domaine(s) d’expertise (section 4.2.3.4).

Dans ce qui suit, nous présentons les principales notations de notre modèle, et plus particulièrement la représentation thématique de la requête, des documents et des collaborateurs sur la base des facettes du besoin en information. Ensuite, nous présentons les deux étapes de calcul des scores et de classification des documents issus de notre approche.

4.2.3.1 Notations : modélisation thématique des entités à partir du besoin à multiples facettes

Cette section introduit les principales notations utilisées dans notre modèle. Plus particulièrement, nous nous concentrons sur la modélisation des entités (requêtes, documents et collaborateurs) qui est différente de celle présentée dans le modèle basé sur les rôles d’expert et de novice de domaine. En effet, alors qu’une simple distribution multinomiale des termes pondérés par un indicateur de spécificité du vocabulaire est suffisante pour décrire l’expertise de collaborateurs caractérisés par une distinction verticale, la modélisation relative à des experts de sous-domaines par rapport à un besoin en information à multiples facettes est plus complexe. En effet, il nous semble important de représenter le degré d’appartenance des entités du modèle avec chacune des facettes du besoin en information. Pour cela, nous favorisons une représentation multi-dimensionnelle où chaque élément représente une facette du besoin en information. Ainsi, chaque entité est représentée par un vecteur de facettes ou sous-thèmes $\omega \in \Omega$ et le poids associé représente l’importance de la facette pour l’entité.

Ainsi, la requête collaborative q sous-jacente au besoin à multiples facettes est modélisée par un vecteur où chaque élément $w_{\omega q}$ représente le poids de la facette ω dans la requête q .

$$q = (w_{1q}, \dots, w_{\omega q}, \dots, w_{\Omega q}) \quad (4.15)$$

De façon similaire, nous représentons les documents d_i (Equation 4.16) et le profil d’expertise $\pi(u_j)^k$ des experts u_j selon le même vecteur thématique correspondant aux facettes du besoin en information (Equation 4.17) :

$$d_i = (w_{1i}, \dots, w_{\omega i}, \dots, w_{\Omega i}) \quad (4.16)$$

$$\pi(u_j)^k = (w_{1j}^k, \dots, w_{\omega j}^k, \dots, w_{\Omega j}^k) \quad (4.17)$$

où $w_{\omega i}$ représente le poids de la facette ω dans le document d_i . $w_{\omega j}^k$ correspond au poids d'expertise de l'utilisateur u_j par rapport à la facette ω à l'itération k . Notons que le profil d'expertise est dynamique compte tenu du processus de réinjection de pertinence dans la mesure où la sélection d'un document par un utilisateur modifie ses connaissances et, par conséquent, son niveau d'expertise de sous-domaines.

4.2.3.2 Identification des facettes du besoin en information partagé

L'objectif de cette étape d'initialisation est d'extraire à partir d'un besoin en information partagé ses facettes afin de déterminer le degré d'expertise vis-à-vis de chacune des facettes auxquelles les utilisateurs peuvent être assimilés dans leur rôle d'expert. Pour cela, nous faisons l'hypothèse que les facettes d'un sujet peuvent être extraites en utilisant l'algorithme du LDA (Blei et al., 2003) à partir d'un ensemble de documents diversifiés.

Dans l'objectif de constituer cet ensemble diversifié, nous nous basons sur la mesure de la pertinence marginale maximale –*Maximal Marginal Relevance*– (Carbonell and Goldstein, 1998), également notée MMR, qui, à partir d'une collection C de documents et d'une requête à multiples facettes q , extrait un ensemble \mathcal{D}^* de n documents qui satisfont à la fois la condition d'être pertinents vis-à-vis de la requête et d'assurer une large diversité thématique. Ainsi, de façon itérative, l'algorithme sélectionne le document d_i avec le score de pertinence RSV_1 et la diversité marginale RSV_2 par rapport à l'ensemble \mathcal{D}^* le plus élevé :

$$\mathcal{D}^* = \arg \max_{d_i \in \mathcal{D}} [\gamma RSV_1(d_i, q) - (1 - \gamma) \max_{d_{i'} \in \mathcal{D}^*} RSV_2(d_i, d_{i'})] \quad (4.18)$$

où $RSV_1(d_i, q)$ exprime les scores de pertinence entre le document d_i et la requête q et $RSV_2(d_i, d_{i'})$ correspond à la similarité de Jaccard entre les deux documents d_i et $d_{i'}$, comme présenté dans la section 4.2.2.1. $\gamma \in [0, 1]$ est un facteur de pondération.

A partir de cet ensemble \mathcal{D}^* de documents diversifiés et de la requête q , nous utilisons l'algorithme du LDA basé sur des probabilités génératives dans l'objectif d'identifier les thèmes latents inclus dans l'ensemble \mathcal{D}^* . Chaque thème ω est assimilé à une facette de la requête q et est associé à deux types de probabilités :

1. $\phi_{t_v|\omega}$ qui estime la probabilité entre un terme t_v et le thème ω .
2. $\theta_{d_i|\omega}$ qui estime la probabilité entre un document d_i et le thème ω .

Le nombre optimal de thèmes ω qui représente les thèmes latents de l'ensemble \mathcal{D}^* est généralement déterminé par maximisation de la vraisemblance, qui est estimée ainsi :

$$l(\mathcal{D}^*|t_v, \omega) = \arg \max_{\Omega} \sum_{t_v \in \mathcal{I}} \log \left(\sum_{\omega \in \Omega} p(t_v|\omega) \right) \quad (4.19)$$

où \mathcal{I} représente l'index extrait de la collection de documents \mathcal{D}^* . La probabilité $p(t_v|\omega)$ correspond à la distribution de probabilité $\phi_{t_v|\omega}$ entre un terme t_v et un thème ω .

L'identification des facettes de la requête collaborative permet d'instancier les représentations thématiques des documents, qui eux-mêmes permettent de modéliser le profil d'expertise des collaborateurs. Dans cet objectif, le poids $w_{\omega i}$ d'un thème ω pour un document d_i est estimé par la probabilité $\theta_{d_i|\omega}$.

Le profil d'expertise $\pi(u_j)^k$ d'un expert u_j est inféré de la distribution thématique extraite à partir des documents diversifiés et des requêtes. Comme source d'évidence du profil d'expertise, nous utilisons les termes des documents sélectionnés $\mathcal{D}(u_j)^k$ par l'expert u_j à l'itération k ainsi que ceux de la requête collaborative q .

Nous proposons un exemple qui repose sur un jeu de données présenté dans l'Illustration 4.7. Dans un premier temps, nous identifions dans l'Illustration 4.8 l'étape de diversification de la collection de documents. Ensuite, nous illustrons la méthodologie pour extraire les facettes et représenter les documents et utilisateurs selon un vecteur thématique dans l'Illustration 4.9.

Illustration 4.7 Description des données illustratives.

Nous considérons une session de RIC avec deux experts, u_1 et u_2 , et qui repose sur une collection C de 20 documents et un index \mathcal{I} de 10 termes.

La distribution des termes dans la collection est la suivante :

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}
d_1	2	1	0	0	0	1	2	0	3	0
d_2	1	0	0	1	1	0	2	0	0	3
d_3	0	0	5	0	5	0	0	2	0	1
d_4	1	2	3	0	1	2	3	2	9	0
d_5	0	1	0	1	0	1	0	1	0	1
d_6	1	0	1	0	0	1	0	1	1	0
d_7	2	0	3	1	0	4	0	2	0	5
d_8	0	1	2	0	2	1	1	2	1	1
d_9	1	1	1	0	3	2	1	0	0	0
d_{10}	0	1	2	0	4	3	2	1	0	0
d_{11}	0	1	3	0	1	0	0	1	2	3
d_{12}	0	0	4	1	1	4	5	6	0	7
d_{13}	3	0	0	1	0	1	3	5	7	9
d_{14}	0	0	3	1	0	0	0	1	0	0
d_{15}	4	2	2	0	2	1	0	0	0	0
d_{16}	0	0	1	1	0	2	0	0	2	1
d_{17}	5	1	0	0	1	0	0	3	2	1
d_{18}	0	2	0	0	1	4	5	0	1	0
d_{19}	1	3	0	0	1	1	1	1	1	1
d_{20}	1	4	0	0	2	0	1	0	1	0

La requête q contient les termes $\{t_1, t_5, t_8, t_{10}\}$.

Illustration 4.8 Construction d'une collection de documents diversifiés.

A partir de la collection C de 20 documents, nous souhaitons extraire le sous-ensemble diversifié \mathcal{D}^* de 10 documents en rapport avec la requête q . Nous présentons dans le tableau les scores des 20 documents pour les 10 itérations nécessaires à la construction du sous-ensemble. Nous posons $\gamma = 0.5$.

Pour l'itération 0, le score représente la similarité entre le document et la requête selon un produit matriciel. Le document d_{13} est le plus pertinent et est donc le premier à être sélectionné pour construire le sous-ensemble.

Pour l'itération 1, le score de chacun des documents est obtenu en combinant leur score de similarité et leur distance par rapport au document d_{13} . Le même calcul est effectué pour les itérations suivantes à la différence que pour le score de diversité, nous considérons la distance maximale entre le document et les documents du sous-ensemble \mathcal{D}^* .

	it0	it1	it2	it3	it4	it5	it6	it7	it8	it9
d1	0.02	0.01	0.01	0.01	0.01	-0.99	-0.99	-0.99	-0.99	-0.99
d2	0.05	0.02	0.02	0.02	0.02	-0.98	-0.98	-0.98	-0.98	-0.98
d3	0.07	0.04	0.03	0.03	0.03	x	x	x	x	x
d4	0.04	0.02	0.02	0.02	0.02	-1.48	-1.48	-1.48	-1.48	-1.48
d5	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	x
d6	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	x	x
d7	0.08	0.04	0.04	0.04	x	x	x	x	x	x
d8	0.05	0.02	0.02	0.02	0.02	-0.48	-0.48	-0.48	-0.48	-0.48
d9	0.04	0.02	0.02	0.02	0.02	-0.48	-0.48	-0.48	-0.48	-0.48
d10	0.05	0.02	0.02	0.02	0.02	-0.98	-0.98	-0.98	-0.98	-0.98
d11	0.05	0.02	0.02	0.02	0.02	0.02	0.02	x	x	x
d12	0.13	0.06	x	x	x	x	x	x	x	x
d13	0.16	x	x	x	x	x	x	x	x	x
d14	0.01	0	0	0	0	0	0	0	0	0
d15	0.06	0.03	0.03	0.02	0.02	x	x	x	x	x
d16	0.01	0	0	0	0	0	0	0	0	0
d17	0.09	0.04	0.04	x	x	x	x	x	x	x
d18	0.01	0	0	0	0	-2.5	-2.5	-2.5	-2.5	-2.5
d19	0.04	0.02	0.02	0.02	0.02	-0.48	-0.48	-0.48	-0.48	-0.48
d20	0.03	0.01	0.01	0.01	0.01	-0.49	-0.49	-0.49	-0.49	-0.49

Au final, $\mathcal{D}^* = \{d_3, d_5, d_6, d_7, d_{11}, d_{12}, d_{13}, d_{15}, d_{16}, d_{17}\}$.

Illustration 4.9 Identification des facettes et représentation thématique.

Nous considérons les notations présentées dans l’Illustration 4.7 et le sous-ensemble de documents \mathcal{D}^* identifié dans l’Illustration 4.8. Nous souhaitons extraire 3 thèmes de la collection, où 3 est choisi arbitrairement pour l’illustration.

Nous faisons l’hypothèse que nous obtenons les distributions de probabilité des termes-thèmes et entités-thèmes, où les entités sont les documents et la requête, suivantes :

	ω_1	ω_2	ω_3		ω_1	ω_2	ω_3
t1	0.15	0.03	0.6	d_3	0.3333	0.3333	0.3333
t2	0.1	0.05	0.02	d_5	0.3333	0.3333	0.3333
t3	0.1	0.05	0	d_6	0.1428	0.1428	0.7142
t4	0.05	0.097	0.02	d_7	0.3333	0.3333	0.3333
t5	0.1	0.003	0.02	d_{11}	0.3333	0.3333	0.3333
t6	0.12	0.05	0.02	d_{12}	0.3333	0.3333	0.3333
t7	0.1	0.05	0.1	d_{13}	0.0526	0.2631	0.6842
t8	0.1	0.07	0.1	d_{15}	0.3333	0.3333	0.3333
t9	0.1	0.05	0.08	d_{16}	0.3333	0.3333	0.3333
t10	0.08	0.55	0.02	d_{17}	0.3333	0.3333	0.3333
				q	0.26	0.32	0.42

La partie de droite du tableau permet d’extraire le vecteur thématique qui représente chaque entité en fonction des 3 thèmes extraits.

Si l’on considère un expert u_1 , son profil d’expertise est initialisé par rapport à la représentation thématique de la requête, soit $\pi(u_j)^0 = (0.26, 0.32, 0.42)$. A l’itération 1, l’expert u_1 a sélectionné le document d_{13} , l’inférence du modèle thématique de la collection permet d’estimer son profil. Nous obtenons par exemple $\pi(u_j)^1 = (0.1, 0.5, 0.5)$.

A la fin de cette étape, les documents $d_i \in \mathcal{D}^*$ sont associés à un score de similarité basé sur la fréquence des termes (Illustration 4.8) et à une représentation thématique. La requête a également une représentation thématique et les profils des deux experts peuvent être inférés des thèmes extraits par le LDA.

Pour l’initialisation des listes de documents retournés aux collaborateurs, nous utilisons une fonction de round-robin. Tout comme dans le chapitre 4, nous fixons le nombre de documents par listes retournées aux utilisateurs à 3, compte tenu de la petite taille de la collection. Ainsi, en classant les documents par score de similarité croissante, nous les distribuons successivement à chacun des utilisateurs (rangs impairs pour u_1 et rangs pairs pour u_2), soit $l^0(u_1, \mathcal{D}^*) = \{d_{13}, d_{17}, d_3\}$ et $l^0(u_2, \mathcal{D}^*) = \{d_{12}, d_7, d_{15}\}$.

4.2.3.3 Calcul des scores de pertinence basé sur les expertises de sous-domaines

Cette étape, présentée dans l’Illustration 4.10, permet d’estimer le score de pertinence des documents en fonction de chacun des experts u_j du groupe de collaboration U selon leur niveau d’expertise par rapport aux facettes de la requête collaborative q . Nous estimons donc la pertinence $p^k(d_i|u_j, q)$ d’un document en fonction d’un expert u_j et de la requête collaborative q à l’itération k ainsi :

$$p^k(d_i|u_j, q) = \frac{p^k(u_j|d_i, q) \cdot p(d_i|q)}{p(u_j|q)} \quad (4.20)$$

Similairement aux hypothèses énoncées dans la section 4.2.2.2, nous considérons que la probabilité $p(u_j|q)$ n’est pas discriminante dans le calcul de la pertinence ainsi que l’expert u_j et la requête q sont indépendants. La probabilité $p^k(d_i|u_j, q)$ peut être ré-écrite comme suit :

$$\begin{aligned} p^k(d_i|u_j, q) &\propto p^k(u_j|d_i, q) \cdot p(d_i|q) \\ &\propto p^k(u_j|d_i) \cdot p(d_i|q) \end{aligned} \quad (4.21)$$

D’une part, la probabilité $p(d_i|q)$ d’un document d_i vis-à-vis de la requête q peut être développée suivant le théorème de Bayes. Nous obtenons la formule suivante :

$$p(d_i|q) = \frac{\frac{p(d_i) \cdot p(q|d_i)}{p(q)}}{\sum_{D_{i'} \in \mathcal{D}^*} \frac{p(d_{i'}) \cdot p(q|d_{i'})}{p(q)}} \propto \frac{p(d_i) \cdot p(q|d_i)}{\sum_{D_{i'} \in \mathcal{D}^*} p(d_{i'}) \cdot p(q|d_{i'})} \quad (4.22)$$

avec \mathcal{D}^* étant la collection de référence. La probabilité $p(d_i)$ d’un document d_i est indépendante de la requête et peut être estimée par un poids uniforme : $p(d_i) = \frac{1}{|\mathcal{D}|}$.

Etant donné la distribution thématique d’un document d_i basée sur les facettes de la requête q , nous estimons la probabilité $p(q|d_i)$ en combinant deux scores de similarité :

1. $RSV_{BM25}(q|d_i)$ estime la pertinence du document d_i vis-à-vis de la requête q selon un modèle probabiliste (Robertson et al., 1995).
2. $RSV_{LDA}(q, d_i)$ estime la pertinence du document d_i vis-à-vis de la requête q selon le modèle d’ordonnancement présenté dans (Harvey et al., 2010) basé sur les distributions de probabilités obtenues par l’algorithme du LDA. Ce score de pertinence est estimé comme suit :

$$RSV_{LDA}(q|d_i) = \prod_{t_v \in q} \sum_{\omega=1}^{\Omega} p(t_v|\omega) \cdot p(\omega|d_i) \quad (4.23)$$

où $p(t_v|\omega)$ représente la probabilité du terme t_v étant donné le thème ω , estimée par $\phi_{t_v|\omega}$. La probabilité $p(\omega|d_i)$ correspond à celle du thème ω sachant le document d_i , précédemment notée $w_{\omega i}$ dans sa représentation thématique.

La probabilité finale $p(q|d_i)$ est ainsi formulée :

$$p(q|d_i) = \lambda RSV_{LDA}(q|d_i) + (1 - \lambda) RSV_{BM25}(q, d_i) \quad (4.24)$$

D'autre part, nous estimons la probabilité $p^k(u_j|d_i)$ par une mesure de similarité de cosinus sim_{cos} , présentée dans la section 2.1.2.2, qui compare les distributions thématiques du document d_i et du profil d'expertise $\pi(u_j)^k$ de l'expert u_j à l'itération k de la façon suivante :

$$p^k(u_j|d_i) = \frac{sim_{cos}(d_i, \pi(u_j)^k)}{\sum_{d_{i'} \in \mathcal{D}^*} sim_{cos}(d_{i'}, \pi(u_j)^k)} \text{ où } sim_{cos}(d_i, \pi(u_j)^k) = \frac{d_i \cdot \pi(u_j)^k}{|d_i| \cdot |\pi(u_j)^k|} \quad (4.25)$$

4.2.3.4 Apprentissage des ordonnancements de documents basé sur les expertises de sous-domaines

Cette étape a pour objectif de maximiser les scores des documents estimés précédemment (Equation 4.20) afin d'assigner un document à l'expert qui est le plus à même de le juger. Nous utilisons l'algorithme d'apprentissage de maximisation de l'espérance mathématique (EM) (Dempster et al., 1977) décomposé en deux étapes "Etape-E" et "Etape-M", comme expliqué dans la section 4.2.2.3. L'objectif de cet algorithme est d'optimiser les probabilités $p(c_{ij} = 1|X_i^k)$ de pertinence $c_{ij} = 1$ d'un document d_i vis-à-vis d'un expert u_j compte tenu des scores X_i^k qu'il a obtenu pour l'ensemble des membres du groupe. Les notations utilisées sont détaillées dans le Tableau 4.2 et l'Algorithme général est illustré dans l'Algorithme 2.

$X_i^k = \{x_{i1}^k, \dots, x_{im}^k\}$	Le vecteur de scores où chaque élément x_{ij}^k est estimé selon l'Equation 4.21 et correspond au score de pertinence d'un document d_i vis-à-vis de chaque expert u_j .
$X^k \in \mathcal{R}^{n \times m}$	La matrice des scores x_{ij}^k pour l'ensemble des documents et des experts.
$c_{ij} = \{0, 1\}$	La variable latente de la non pertinence ou la pertinence d'un document d_i pour un expert u_j .
ϕ_j^k	La fonction de densité de la loi gaussienne modélisant la pertinence des documents par rapport à l'expert u_j à l'itération k .
θ_j^k	Les paramètres des distributions des scores à l'itération k étant donné l'expert u_j correspondant à la loi gaussienne ϕ_j^k , à savoir μ_j^k et σ_j^k .
α_j^k	Le coefficient du modèle de mélange.

TABLE 4.2: Notations utilisées dans l'algorithme d'allocation des documents aux experts de sous-domaines

Illustration 4.10 Estimation de la pertinence des documents.

Nous considérons les notations présentées dans l'illustration 4.7 et les calculs effectués dans les exemples précédents. Faisons l'hypothèse que l'expert u_1 sélectionne à l'itération 1 le document d_3 . Notre objectif est de calculer un score de pertinence pour chaque document non sélectionné $\mathcal{D}_{ns}^1 = \{d_3, d_5, d_6, d_7, d_{11}, d_{12}, d_{15}, d_{16}, d_{17}\}$ pour chacun des deux experts. Nous déroulons les calculs pour le document d_3 uniquement.

L'expert u_1 a sélectionné le document d_3 et est caractérisé par le profil d'expertise $\pi(u_1)^1 = (0.1, 0.5, 0.4)$. Le score $p^1(d_3|u_1, q)$, sans pratiquer les normalisations, est égal à $0.1 * 0.1 * 0.84 = 0,0084$ où :

- Le score $p(q|d_3) = 0.12 * 0,5 + 0.07 * 0.5 = 0.1$ est une combinaison linéaire, avec $\lambda = 0.5$, de $RSV_{BM25}(q|d_3) = 0.07$ que nous estimons par le score calculé dans l'exemple 4.8 et $RSV_{LDA}(q, d_3) = 0.12$ estimé à partir des distributions terme-thème et document-thème selon l'équation 4.23.
- Le score $p^k(u_1|d_3) = \frac{0.33*0.1+0.33*0.5+0.33*0.4}{0.65*0.58} = 0.89$ est estimé par une mesure de cosinus entre la représentation thématique du document d_3 et de l'expert u_1 .

L'expert u_2 n'a pas sélectionné de document et est caractérisé par le profil d'expertise $\pi(u_2)^1 = (0.26, 0.32, 0.42)$. Le score $p^1(d_3|u_2, q)$, sans pratiquer les normalisations, est égal à $0.1 * 0.1 * 1 = 0.01$ où :

- Le score $p(q|d_3) = 0.12 * 0,5 + 0.07 * 0.5 = 0.01$ est une combinaison linéaire, avec $\lambda = 0.5$, de $RSV_{BM25}(q|d_3) = 0.12$ que nous estimons par le score calculé dans l'exemple 4.8 et $RSV_{LDA}(q, d_3) = 0.07$ estimé à partir des distributions terme-thème et document-thème selon l'équation 4.23.
- Le score $p^k(u_2|d_3) = \frac{0.33*0.26+0.33*0.32+0.33*0.42}{0.59*0.58} = 0.98$ est estimé par une mesure de cosinus entre la représentation thématique du document d_3 et de l'expert u_2 .

Les scores finaux incluant les différentes normalisations, arrondis à 10^{-2} , sont recensés dans le tableau suivant :

d_i	$p^1(d_i u_1, q)$	$p^1(d_i u_2, q)$
d_3	0.14	0.14
d_5	0.1	0.09
d_6	0.02	0.02
d_7	0.15	0.15
d_{11}	0.12	0.12
d_{12}	0.1	0.09
d_{15}	0.13	0.13
d_{16}	0.09	0.1
d_{17}	0.15	0.16

Algorithm 2: Algorithme d'ordonnancement collaboratif basé sur la maximisation de l'espérance.

Données: $\mathcal{D}_{ns}^k, X^k, U, u_j$

Résultat: $l^k(u_j, \mathcal{D}_{ns}^k)$

Début

$l^k(u_j, \mathcal{D}_{ns}^k) = \{\}$

/* Etape 1: Apprentissage de l'appariement document-collaborateur ***/**

Répéter

/* Etape-E ***/**

Pour chaque document $d_i \in \mathcal{D}_{ns}^k$ **faire**

Pour chaque expert $e_{j'} \in U$ **faire**

$$\begin{aligned} p(c_{ij'} = 1 | X_i^k) &= \frac{\alpha_{j'} \phi_j(x_{ij'}^k)}{\sum_{l=1}^m \alpha_l \phi_l(x_{il}^k)} \\ M_{ij}^{EM,k} &= p(c_{ij'} = 1 | X_i^k) \end{aligned}$$

/* Etape-M ***/**

$$S_{j'}^k = \sum_{d_h \in \mathcal{D}_{ns}^k} p(c_{ij'} = 1 | X_h^k)$$

$$\alpha_{j'}^k = \frac{1}{n} S_{j'}^k$$

$$\mu_{j'}^k = \frac{1}{S_{j'}^k} \sum_{d_h \in \mathcal{D}_{ns}^k} p(c_{ij'} = 1 | X_h^k) \cdot x_{hj'}^k$$

$$\sigma_{j'}^k = \frac{1}{S_{j'}^k} \sum_{d_h \in \mathcal{D}_{ns}^k} p(c_{ij'} = 1 | X_h^k) \cdot (x_{hj'}^k - \mu_j)^2$$

jusqu'à convergence

/* Etape 2: Allocation des documents aux utilisateurs ***/**

$$M^{EM,k} = \text{odds}(M^{EM,k})$$

Pour chaque document $d_i \in \mathcal{D}_{ns}^k$ **faire**

Si $\text{argmax}_{ij} M_{ij}^{EM,k} = u_j$ **alors**

$$l^k(u_j, \mathcal{D}_{ns}^k) = l^k(u_j, \mathcal{D}_{ns}^k) \cup d_i$$

Retourner $l^k(u_j, \mathcal{D}_{ns}^k)$

1. **Apprentissage de l'appariement document-collaborateur.** L'objectif, ici est d'optimiser les scores des documents et repose sur l'algorithme de maximisation de l'espérance EM selon deux étapes (Illustration 4.11) :

- *Etape-E.* Nous estimons la probabilité $p(c_{ij} = 1 | X_i^k)$ de pertinence du document d_i vis-à-vis de l'expert u_j à l'iteration k compte tenu du vecteur de score X_i^k du document d_i ainsi que de l'ensemble des experts du groupe de collaboration. Cette probabilité $p(c_{ij} | X_i^k)$ est estimée par un modèle à mélange qui repose sur des lois de probabilité gaussienne ϕ_j et qui modélise la pertinence c_{ij} des documents pour l'expert u_j à l'iteration k :

$$p(c_{ij} = 1 | X_i^k) = \frac{p(c_{ij} = 1)p(x_{ij}^k | c_{ij} = 1)}{p(c_{ij} = 1)p(x_{ij}^k | c_{ij} = 1) + p(c_{ij} = 0)p(x_{ij}^k | c_{ij} = 0)}$$

(4.26)

Dans cette équation, nous faisons l'hypothèse que la non pertinence, notée $c_{ij} = 0$, d'un document d_i vis-à-vis d'un expert u_j peut être formulée par la probabilité de pertinence vis-à-vis d'un autre expert du groupe, notée $c_l = 1 \forall l = \{1, \dots, m\}$ with $l \neq j$. Ainsi, le dénominateur correspond à la somme des probabilités de pertinence, notée $c_l = 1$, par rapport aux experts e_l . par conséquent, nous remplaçons les probabilités p_{ij}^k et p_{il}^k par des lois gaussiennes $\phi_j(x_{ij}^k)$ et $\phi_l(x_{il}^k)$ afin d'obtenir l'estimation finale de la probabilité.

$$p(c_{ij} = 1 | X_i^k) \propto \frac{\alpha_j p_{ij}^k}{\alpha_j p_{ij}^k + (1 - \alpha_j) \bar{p}_{ij}} = \frac{\alpha_j p_{ij}^k}{\sum_{l=1}^m \alpha_l p_{il}^k} \quad (4.27)$$

$$\text{avec } \begin{cases} p_{ij}^k = p(c_{ij} = 1) p(x_{ij}^k | c_{ij} = 1) \\ \bar{p}_{ij}^k = p(c_{ij} = 0) p(x_{ij}^k | c_{ij} = 0) \end{cases}$$

— *Etape-M.* Cette étape met à jour les paramètres θ_j et maximise la log-vraisemblance des données complétées $\mathcal{L}(R_j = Rel | X_j^k, \theta_j^k)$, estimée de la façon suivante :

$$\mathcal{L}(c_{ij} = 1 | X_j^k, \theta_j) = \sum_{h=1}^n \sum_{l=1}^m \log(p(c_{hl} = 1 | X_h^k)) p(c_{hl} = 1 | X_h^k) \quad (4.28)$$

2. **Allocation des documents aux collaborateurs.** Une fois les scores document-expert optimisés, l'enjeu est de déterminer quel expert est le plus à même de juger un document. Similairement à l'hypothèse probabiliste, nous déterminons pour chaque élément de la matrice $M_{ij}^{EM,k}$ une mesure appelée *odds value* et notée $odds(M_{ij}^{EM,k})$, estimée par le ratio entre la probabilité de pertinence $M_{ij}^{EM,k}$ et la probabilité de non pertinence, estimée par $\sum_{l=1, l \neq j}^m M_{il}^{EM,k}$. La classification des documents est réalisée sur la base des éléments de la matrice $M^{EM,k}$ mis à jour selon cette mesure :

$$\forall d_i, \exists u_j^*; u_j^* = \arg \max_{u_j \in U} M_{ij}^{EM,k} \quad (4.29)$$

Egalement, nous considérons le paradigme de la division du travail en assignant à un expert seulement les documents qui ne sont pas simultanément retournés aux autres membres du groupe de collaboration.

Illustration 4.11 Apprentissage de l'ordonnancement collaboratif des documents.

Nous considérons les calculs effectués dans l'illustration 4.10 et souhaitons appliquer l'algorithme EM afin de maximiser les scores. Nous présentons seulement les 2 premières itérations.

Pour l'optimisation du score du document d_3 vis-à-vis de l'expert u_1 à l'itération 1, par exemple, $x_{31}^1 = \frac{0.11 \cdot \mathcal{N}(0.141, 0.11, 0.04)}{0.11 \cdot \mathcal{N}(0.141, 0.11, 0.04) + 0.89 \cdot \mathcal{N}(0.143, 0.11, 0.05)} = 0.114$.

Pour l'itération 2, $x_{31}^2 = \frac{0.11 \cdot \mathcal{N}(0.114, 0.11, 0.001)}{0.11 \cdot \mathcal{N}(0.114, 0.11, 0.001) + 0.89 \cdot \mathcal{N}(0.108, 0.10, 0.002)} = 0.184$.

Les résultats finaux et normalisés sont synthétisés dans le tableau suivant :

		Itération k=0		Itération k=1		Itération k=2	
		x_{i1}^0	x_{i2}^0	x_{i1}^1	x_{i2}^1	x_{i1}^2	x_{i2}^2
Etape-E	d_3	0,141	0,143	0,114	0,108	0,184	0,003
	d_5	0,101	0,092	0,118	0,103	0,258	0
	d_6	0,017	0,02	0,087	0,14	0	0,464
	d_7	0,147	0,15	0,114	0,108	0,219	0,002
	d_{11}	0,121	0,123	0,112	0,11	0,029	0,053
	d_{12}	0,097	0,089	0,118	0,104	0	0
	d_{15}	0,128	0,13	0,112	0,109	0,072	0,019
	d_{16}	0,094	0,096	0,109	0,113	0	0,457
	d_{17}	0,154	0,157	0,115	0,107	0,239	0,001
Etape-M	μ_i^k	0.11	0.11	0.11	0.10	$9.47 \cdot 10^{-5}$	0.0004
	σ_j^k	0.04	0.05	0.001	0.002	$4.40 \cdot 10^{-9}$	$2.63 \cdot 10^{-7}$
	α_j^k	0.11	0.11	0.11	0.11	0.40	0.24
	vraiss.			-1.88		-0.73	

Nous faisons l'hypothèse que l'algorithme EM a atteint sa convergence à l'itération 2 et nous souhaitons maintenant établir la liste de l'expert u_1 . Tout d'abord, nous extrayons les documents d_i pour lesquels le score optimisé x_{i1}^2 vis-à-vis de l'expert u_1 est supérieur à celui de l'expert u_2 , noté x_{i2}^2 . A l'issue de l'algorithme EM, nous retenons les documents d_3 , d_5 , d_7 , d_{15} et d_{17} comme candidats à la liste retournée à l'utilisateur. Selon le paradigme de la division du travail, nous considérons seulement les documents qui ne sont pas dans la liste retournée à l'expert u_2 , soit les documents d_3 , d_5 et d_{17} . En les triant par score décroissant, nous construisons ainsi la liste $l^1(u_1, \mathcal{D}_{ns}^1) = \{d_5, d_{17}, d_3\}$.

4.3 Expérimentations et résultats

Afin de valider nos deux contributions, nous avons mis en place un protocole commun basé sur la simulation que nous présentons dans ce qui suit. Nous soulignons le fait que certaines caractéristiques du protocole sont inhérentes aux spécificités de nos contributions et que, par conséquent, nous détaillons les

procédures mises en place pour chacune des contributions. Par la suite, nous détaillons les résultats qui ont permis de tester l'efficacité de chacune de nos contributions.

4.3.1 Cadre expérimental

A notre connaissance, hormis les collections propriétaires (Morris et al., 2006; Shah and González-Ibáñez, 2011b), il n'existe pas à ce jour de collection de test standard pour l'évaluation de l'efficacité de la RIC. Pour répondre à ce besoin, Foley and Smeaton (2009) ont proposé un protocole de simulation de la collaboration qui exploite les fichiers log de la collection TREC Interactive 6-7-8¹. La simulation présente de nombreux avantages par rapport aux expérimentations utilisateurs. Entre autres, elle offre la possibilité d'évaluer plusieurs scénarios ou d'effectuer une étape de paramétrage, qui seraient bien trop coûteux en temps et en besoin humain dans le cadre d'une expérimentation avec des utilisateurs réels (White et al., 2005). Pour ces raisons, nous proposons de suivre ce cadre expérimental afin de valider nos deux contributions.

4.3.1.1 Données expérimentales

Collection de documents. Notre expérimentation repose sur la campagne d'évaluation TREC 6-7-8 Interactive, précédemment utilisée dans un cadre d'évaluation de RIC (Foley, 2008; Foley and Smeaton, 2009). Cette collection est adaptée à la résolution de tâches complexes dans la mesure où un des objectifs des participants est d'identifier les différents aspects thématiques reliés au besoin en information initial (Over, 2001). La collection utilisée est TREC Financial Times of London 1991-1994 (disque 4 de la Collection TREC ad-hoc) qui inclut 210 158 articles, 20 sujets de recherche (*topics*).

La tâche de la campagne TREC Interactive imposait aux participants de fournir deux types de fichiers log : (a) des fichiers logs clairsemés – "*sparse format data*" – qui incluent la liste des documents sélectionnés par les participants ainsi que (b) des fichiers logs exhaustifs – "*rich format data*" – qui englobent de nombreuses méta-données supplémentaires, telles que les requêtes soumises, les documents visités et/ou sélectionnés et les indicateurs temporels des différentes actions. Ces indicateurs temporels expriment en secondes le temps écoulé depuis le début de la session jusqu'à ce que l'utilisateur réalise l'action. C'est ce dernier type de fichiers log – "*rich format data*" –, illustré dans la Figure 3.12 de la section 3.6.2.1, que nous avons considéré dans notre expérimentation.

Dans le cas du modèle basé sur les rôles d'expert de sous-domaines, nous avons vérifié que les sujets TREC correspondent à des thèmes exploratoires et à multiples-facettes en estimant la distance de Jaccard paire à paire entre les 1000 documents les plus pertinents. Pour chacun des sujets, nous avons évalué

1. <http://trec.nist.gov>

la couverture de la diversité à une valeur proche de 1, ce qui nous permet de confirmer que ces sujets sont exploratoires, comme mentionné dans les configurations de la Collection TREC Interactive.

Pour chaque sujet de recherche TREC, l'ensemble des documents pertinents (vérité de terrain) –*groundtruth*– est formé selon un niveau d'accord –*agreement level*– entre les participants en s'assurant qu'au moins deux utilisateurs ont sélectionné un document pour un même sujet.

Simulation de la collaboration. Nous avons étendu le cadre expérimental présenté par Foley and Smeaton (2009) selon deux aspects : (a) la création des groupes de collaboration basés sur l'expertise de domaine et (b) l'adaptation des mesures d'efficacité de la session au processus itératif de notre modèle et aux rôles utilisateurs. Dans ce qui suit, nous analysons les principaux composants de notre protocole d'évaluation :

1. *Les collaborateurs.* Compte tenu des caractéristiques de nos deux contributions (collaboration réalisée par une paire d'utilisateurs dans le cadre du modèle basé sur les rôles d'expert et de novice de domaine ou un groupe de taille plus large dans le cas du modèle basé sur les experts de sous-domaines), nous avons effectué deux protocoles différents pour chacun de nos modèles afin de construire les groupes de collaboration. Nous les présentons ci-dessous.
 - (a) *Modèle basé sur les rôles d'expert et de novice.* Dans notre scénario, la collaboration requiert une paire $c(u_j, u_{j'})$ d'utilisateurs u_j et $u_{j'}$ qui partagent le même besoin en information. Cette paire, est identifiée au travers des fichiers log fournis par l'ensemble des participants de la collection TREC Interactive. Une fois le niveau d'expertise $Expertise(u_j, \tau)$ identifié pour chaque utilisateur u_j pour un sujet TREC τ , Nous avons considéré deux méthodes de construction des paires de collaborateurs, dont un exemple est présenté dans l'illustration 4.12 :
 - i. *La méthode exhaustive*, notée *ME_{ex}*, considère toutes les combinaisons possibles $c(u_j, u_{j'})$ d'utilisateurs u_j et $u_{j'}$, appartenant au même groupe de participants et ayant réalisé une session reliée au même sujet TREC τ . Les rôles sont ensuite attribués aux utilisateurs selon le principe de différence relative de l'expertise : l'utilisateur u_j avec le niveau d'expertise $Expertise(u_j, \tau)$ le plus élevé reçoit le rôle d'expert du sujet de la requête et l'autre utilisateur $u_{j'}$ est considéré comme novice. Nous avons ainsi dénombré 243 paires de collaborateurs, et par conséquent 243 requêtes et sessions collaboratives.
 - ii. *La méthode sélective*, notée *MS_{el}*, construit les paires à partir des collaborateurs ayant été détectés comme ayant un niveau d'expertise aux extrêmes du spectre. Dans cet objectif, nous avons effectué une classification, en utilisant l'algorithme des 2-means des utilisateurs ayant interagi avec le système sur une session reliée au même

sujet TREC τ ; nous considérons comme critère de classification le niveau d'expertise $Expertise(u_j, \tau)$ de chaque utilisateur u_j . Nous obtenons ainsi deux classes d'utilisateurs pour chaque sujet TREC τ , respectivement une pour les experts et une autre pour les novices. Ensuite, nous avons réalisé toutes les combinaisons possibles $c(u_j, u_{j'})$ d'expert de domaine u_j et de novice du domaine $u_{j'}$, identifiés lors de la classification. Ainsi, nous avons identifié jusqu'à 95 paires de collaborateurs, respectivement associées à 95 requêtes et sessions collaboratives.

- (b) *Modèle basé sur les rôles d'expert de sous-domaines.* Similairement à la méthode utilisée pour la détection des experts et des novices, pour chaque sujet TREC τ , nous avons effectué une classification des utilisateurs ayant participé à la résolution du sujet. Pour cela, nous avons utilisé l'algorithme des 2-means sur le critère d'expertise des utilisateurs. Nous obtenons ainsi deux classes utilisateurs que nous avons labellisées comme des classes d'experts et de novices selon la valeur moyenne obtenue par la classe. Nous avons ainsi combiné les utilisateurs de la classe "expert" pour former des groupes de taille $m \geq 2$. Nous avons obtenu 277 groupes d'experts sur 19 sujets TREC pouvant varier de 2 à 6 collaborateurs, comme illustré dans le Tableau 4.3. Notons que le sujet TREC 438i ne comporte qu'un seul expert, ce qui ne permet donc pas de construire des groupes de collaborateurs.

Taille du groupe	2	3	4	5	6	Total
Nombre de groupes	129	88	44	14	2	277

TABLE 4.3: Répartition des groupes d'experts selon leur effectif

2. *Les rôles.* Compte tenu que nos contributions sont basées sur des rôles d'expertise, l'attribution de ces rôles aux collaborateurs impliqués dans la paire $c(u_j, u_{j'})$ doit être réalisée en fonction des hypothèses formulées par Allen (1991) et White and Dumais (2009). Nous proposons d'estimer le niveau d'expertise $Expertise(u_j, \tau)$ d'un utilisateur u_j pour le sujet TREC τ selon deux indicateurs basés sur l'analyse du vocabulaire des utilisateurs, comme illustré dans l'Exemple 4.14 :

- (a) *L'expertise basée sur la spécificité*, notée $Expertise_S(u_j, \tau)$, dépend de la spécificité moyenne $Spec(d_i)$, détaillée dans l'Equation 4.1, des documents sélectionnés $d_i \in \mathcal{D}^{\delta_\tau}(u_j)$ par l'utilisateur u_j dans une session de recherche δ_τ appartenant au sujet τ :

$$Expertise_S(u_j, \tau) = \frac{\sum_{d_i \in \mathcal{D}^{\delta_\tau}(u_j)} Spec(d_i)}{|\mathcal{D}^{\delta_\tau}(u_j)|} \quad (4.30)$$

avec $|\mathcal{D}^{\delta_\tau}(u_j)|$ correspondant au nombre de documents sélectionnés par l'utilisateur u_j lors de la session δ_τ .

Illustration 4.12 Méthode de construction des paires d'expert et de novice.

Nous souhaitons construire des paires de collaborateurs à partir de 5 utilisateurs associés à un niveau d'expertise :

$(u_1, 0.3), (u_2, 0.1), (u_3, 0.9), (u_4, 0.56), (u_5, 0.87)$.

Méthode exhaustive : En combinant tous les utilisateurs, nous obtenons les 10 paires suivantes :

$(u_1 \text{ "expert" }, u_2 \text{ "novice" }), (u_1 \text{ "novice" }, u_3 \text{ "expert" }),$
 $(u_1 \text{ "novice" }, u_4 \text{ "expert" }), (u_1 \text{ "novice" }, u_5 \text{ "expert" }),$
 $(u_2 \text{ "novice" }, u_3 \text{ "expert" }), (u_2 \text{ "novice" }, u_4 \text{ "expert" }),$
 $(u_2 \text{ "novice" }, u_5 \text{ "expert" }), (u_3 \text{ "expert" }, u_4 \text{ "novice" }),$
 $(u_3 \text{ "expert" }, u_5 \text{ "novice" }), (u_4 \text{ "novice" }, u_5 \text{ "expert" }).$

Méthode sélective : En pratiquant une 2-means, nous obtenons 2 classes basées sur les rôles utilisateur :

$Expert = \{u_3, u_4, u_5\}$ et $Novice = \{u_1, u_2\}$.

En combinant un utilisateur de la classe "Expert" avec un de la classe "Novice", nous obtenons les 6 paires suivantes :

$(u_3 \text{ "expert" }, u_1 \text{ "novice" }), (u_3 \text{ "expert" }, u_2 \text{ "novice" }),$
 $(u_4 \text{ "expert" }, u_1 \text{ "novice" }), (u_4 \text{ "expert" }, u_2 \text{ "novice" }),$
 $(u_5 \text{ "expert" }, u_1 \text{ "novice" }), (u_5 \text{ "expert" }, u_2 \text{ "novice" }).$

Illustration 4.13 Méthode de construction des groupes d'experts.

A partir des données présentées dans l'illustration 4.12, nous pouvons construire les groupes de collaborations suivants :

$(u_3, u_4), (u_3, u_5), (u_4, u_5)$ et (u_3, u_4, u_5) .

- (b) *L'expertise basée sur l'autorité*, notée $Expertise_A(u_j, \tau)$, qui repose sur le score d'autorité présenté dans les travaux de Foley and Smeaton (2009). L'expertise basée sur l'autorité $Expertise_A(u_j, \tau)$ pour un utilisateur u_j par rapport au sujet τ dépend des termes inclus dans les documents pertinents identifiés pour le sujet TREC τ et dans les documents $d_i \in \mathcal{D}^{\delta_\tau}(u_j)$ sélectionnés par l'utilisateur u_j pendant la session de recherche interactive δ_τ connectée au sujet τ . Cette mesure est estimée par la corrélation entre les poids des termes des documents appartenant aux deux catégories énoncées précédemment.

Illustration 4.14 Estimation du niveau d'expertise des collaborateurs.

Nous considérons la collection de documents présentée dans l'Illustration 4.1. Soit un utilisateur u_2 ayant sélectionné les documents d_4 et d_5 . Nous souhaitons estimer son niveau d'expertise par rapport à un sujet τ .

Mesure basée sur la spécificité : En tenant compte des niveaux de spécificité estimés dans l'Illustration 4.2, nous pouvons en déduire que le niveau d'expertise de l'utilisateur u_2 par rapport à un sujet τ est égal à $Expertise_S(u_2, \tau) = \frac{0.15+0.1}{2} = 0.125$.

Mesure basée sur l'autorité : A partir des documents d_4 et d_8 , nous pouvons construire un vecteur moyen de poids $d' = (0.525, 0.225, 0.125, 0.125)$. Faisons l'hypothèse que les documents d_1 , d_4 et d_9 sont pertinents pour le sujet τ . Le vecteur moyen des poids des documents inclus dans l'oracle est $o = (0.36, 0.34, 0.21, 0.05)$. Le niveau d'expertise de l'utilisateur u_2 par rapport à un sujet τ est égal à $Expertise_A(u_2, \tau) = \rho(d', o) = 0.77$.

3. *Le besoin en information partagé*. L'objectif d'une session de RIC consiste en la résolution d'un besoin en information partagé (Foley and Smeaton, 2010). Pour extraire le besoin en information partagé, nous considérons l'hypothèse exprimée par Foley (2008) qui repose sur le fait que les collaborateurs forment ensemble la requête initiale pour un besoin en information partagé. Cette hypothèse est renforcée par le fait que, dans la collection TREC Interactive, les requêtes incluent généralement 3 termes et 94% d'entre elles ont au moins un mot en commun (Foley, 2008). Compte tenu des spécificités de nos contributions, nous proposons deux méthodes de construction du besoin en information :

- (a) *Modèle basé sur les rôles d'expert et de novice*. Afin d'identifier le besoin en information partagé entre les collaborateurs, nous exploitons les fichiers log des participants pour construire, pour chaque paire $c(u_j, u_{j'})$ de collaborateurs, la requête collaborative q à partir de la fusion des requêtes soumises par les collaborateurs u_j et $u_{j'}$, comme présenté dans l'Exemple 4.15. Respectivement, pour chaque méthode de construction des paires exhaustives et sélectives, nous avons au total 243 et 95 requêtes collaboratives. En résumé, pour un sujet TREC τ et une paire $c(u_j, u_{j'})$ de collaborateurs, une unique requête collaborative q est générée. Nous soulignons le fait que plusieurs requêtes collaboratives $q \in Q_\tau$ sont assignées à un sujet TREC τ , respectivement associées à un groupe de collaboration.

Illustration 4.15 Estimation du niveau d'expertise des collaborateurs.

Pour le sujet TREC 326i, si deux utilisateurs d'une même paire ont respectivement soumis les requêtes "ferry sinking death" et "ferry sink people", alors la requête collaborative du groupe est la suivante : "ferry sink death people".

- (b) *Modèle basé sur les rôles d'expert de sous-domaine.* Afin de modéliser le besoin en information par une requête collaborative, nous considérons le sujet TREC, à l'inverse du modèle de RIC basé sur les rôles d'expert et de novice où nous avons considéré les requêtes reformulées par les collaborateurs. Notre choix est justifié par le fait que les requêtes reformulées constituent une perception individuelle du besoin en information et traduisent un seul aspect du besoin en information en fonction des connaissances de l'utilisateur. Ainsi, nous considérons que le sujet TREC constitue une requête plus objective et qui couvre l'ensemble des facettes du besoin en information partagé.

4. *La session de recherche collaborative.* Nous avons simulé les sessions de RIC à partir des fichiers log des sessions individuelles. Plus particulièrement, nous les avons fusionnés afin d'obtenir la chronologie des documents sélectionnés lors de la session de collaboration par les deux utilisateurs u_j et $u_{j'}$. Cette synchronisation des sessions de recherche est illustrée dans la figure 4.4. Notons que pour les besoins de notre expérimentation, seule la synchronisation des documents sélectionnés est nécessaire car notre modèle tient compte d'une unique requête collaborative, et non de requêtes reformulées individuellement au cours de la session. Plus formellement, nous obtenons une liste $L^S(u_j, u_{j'})$ des documents sélectionnés $\mathcal{D}^S(u_j, u_{j'})$ par la fusion chronologique des listes $l^S(u_j)$ and $l^S(u_{j'})$ qui englobent respectivement les documents sélectionnés par les utilisateurs u_j et $u_{j'}$. La méthodologie globale est illustrée dans l'Illustration 4.16.

Lors de la session, un document inclu dans la liste synchronisée $L^S(u_j, u_{j'})$ doit satisfaire deux conditions pour déclencher une itération k de notre modèle, qui consiste à mettre à jour la liste $l^k(u_j, \mathcal{D}_{ns}^k)$ des documents non sélectionnés \mathcal{D}_{ns}^k :

- (a) Etre inclu dans les listes $l^k(u_j, \mathcal{D}_{ns}^k)$ des documents couramment affichés à chaque collaborateur u_j .
- (b) Ne pas avoir déjà été sélectionné par l'autre collaborateur $u_{j'}$ dans l'objectif de garantir le paradigme de la division du travail.

De plus, dans l'objectif de rendre l'expérimentation la plus naturelle possible et similairement à Foley and Smeaton (2009), nous considérons un niveau maximum de rappel à 30 compte tenu que les utilisateurs évaluent généralement autour de 30 documents dans une liste. Par conséquent, nous prenons seulement en compte les documents sélectionnés, appelés aussi jugements de pertinence, exprimés par rapport au 30 premiers documents de la liste retournée à l'utilisateur dans l'objectif de déclencher une itération de notre

modèle. Ainsi, la liste synchronisée $L^S(u_j, u_{j'})$ des documents sélectionnés est considérée dynamiquement par rapport aux listes $l^k(u_j, \mathcal{D}_{ns}^k)$ retournées à chaque collaborateur u_j pour une itération k . Nous soulignons qu'une session S implique une paire $c(u_j, u_{j'})$ de collaborateurs u_j et $u_{j'}$, un sujet TREC τ et une requête collaborative q . De plus, pour chaque sujet TREC τ , plusieurs sessions collaboratives δ_τ peuvent être simulées.

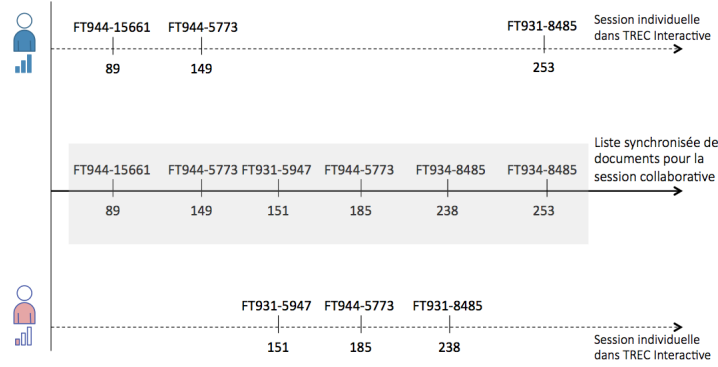


FIGURE 4.4: Simulation d'une session de RIC à partir des fichiers log fournis par le participant rmitMG6 du sujet TREC 326i

Illustration 4.16 Simulation d'une session de RI collaborative.

Soit un ensemble de fichiers log qui caractérise les actions suivantes, comme illustré dans la Figure 4.4 :

- L'utilisateur u_1 a sélectionné trois documents : FT944-15661, FT944-5773 et FT931-8485, respectivement 89, 149 et 253 secondes après le début de la session.
- L'utilisateur u_2 a sélectionné trois documents : FT931-5947, FT944-5773 et FT931-8485, respectivement 151, 185 et 238 secondes après le début de la session.

Nous pouvons ainsi extraire la liste synchronisée suivante :
 $L^S(u_1, u_2) = \{(FT944-15661, u_1), (FT944-5773, u_1), (FT931-5947, u_2), (FT944-5773, u_2), (FT931-8485, u_2), (FT931-8485, u_1)\}.$

Ainsi, la première itération est lancée seulement si le document FT944-15661 appartient à la liste $l^1(u_j, \mathcal{D}_{ns}^1)$ des documents retournés à l'utilisateur u_1 . Un peu plus tard dans la session, l'itération 4 est lancée à la condition que le document FT944-5773 appartienne à la liste $l^4(u_j, \mathcal{D}_{ns}^4)$ de l'utilisateur u_2 et que l'itération 2 n'ait pas été lancée, ce qui serait contraire au paradigme de la division du travail car le document aurait déjà été sélectionné par l'utilisateur u_1 .

4.3.1.2 Métriques d'évaluation

Conformément aux objectifs de notre évaluation, nous distinguons deux grandes catégories de mesures. Similairement à Foley (2008), ces mesures sont appliquées au rang 30 et sont notées $P@30$, $P@30^r$, $P@30^k$, $C@30$ et $PC@30$ dans ce qui suit.

1. *Les mesures orientées précision* : elles permettent de mesurer l'efficacité en termes de nombre de documents pertinents retournés. Dans l'objectif d'éviter le biais induit par des mesures classiques qui favorisent les sessions de recherche longues et compte tenu du contexte interactif des sessions de RIC, nous proposons de considérer une mesure moyenne d'agrégation des précisions au niveau de la session, appelée la micro-précision (Sebastiani, 2002). Etant donné ces hypothèses, nous formalisons trois niveaux d'analyse :

- (a) La micro-précision moyenne au niveau de la session P_{micro} estime la précision moyenne des sessions de recherche en faisant abstraction des ordonnancements intermédiaires retournés au cours de la session :

$$P_{micro} = \frac{1}{|\theta|} \sum_{\tau \in \theta} \frac{1}{|\delta_\tau|} \sum_{S \in \delta_\tau} \frac{\sum_{l \in \{L_S\}} Pert_l}{\sum_{l \in \{L_S\}} Sel_l} \quad (4.31)$$

où θ est l'ensemble des sujets TREC et δ_τ représente l'ensemble des sessions de recherche S reliées au sujet TREC τ . L_S correspond à l'ensemble des listes ordonnées présentées aux collaborateurs dans la session de recherche S . $Pert_l$ et Sel_l expriment le nombre de documents pertinents, respectivement sélectionnés, dans la liste l .

- (b) La micro-précision moyenne au niveau des rôles $P_{micro}^{r,S}$ estime la précision moyenne des sessions de recherche en prenant en compte seulement les utilisateurs caractérisés par le rôle r :

$$P_{micro}^{r,S} = \frac{1}{|\theta|} \sum_{\tau \in \theta} \frac{1}{|\delta_\tau|} \sum_{S \in \delta_\tau} \frac{\sum_{l \in \{L_{S,r}\}} Pert_l}{\sum_{l \in \{L_{S,r}\}} Sel_l} \quad (4.32)$$

où $L_{S,r}$ représente l'ensemble des ordonnancements présentés aux collaborateurs caractérisés par le rôle r dans la session de recherche S . Le rôle de l'utilisateur est noté $r \in \mathcal{R} = \{\text{expert}, \text{novice}\}$.

- (c) La micro-précision moyenne au niveau de l'itération P_{micro}^k estime la précision moyenne des listes de documents retournées aux deux utilisateurs à l'itération k :

$$P_{micro}^k = \frac{1}{|\theta|} \sum_{\tau \in \theta} \frac{1}{|\delta_\tau|} \sum_{S \in \delta_\tau} \frac{\sum_{l \in \{L_{S,k}\}} Pert_l}{\sum_{l \in \{L_{S,k}\}} Sel_l} \quad (4.33)$$

où $L_{S,k}$ représente l'ensemble des ordonnancements présentés aux collaborateurs dans la session S à l'itération k .

2. *Les mesures orientées collaboration* : elles permettent de mesurer l'efficacité en terme de diversité des documents retournés aux deux utilisateurs, sur l'ensemble de la session de recherche. A cet effet, nous considérons deux indicateurs basés sur la mesure de couverture, décrite dans Shah (2012) et adaptée au contexte d'évaluation des ordonnancements des documents :

- (a) Le ratio de couverture Co analyse la diversité moyenne des listes de documents retournées pendant les sessions de recherche :

$$Co = \frac{1}{|\theta|} \sum_{\tau \in \theta} \frac{1}{|\delta_\tau|} \sum_{S \in \delta_\tau} \frac{Couverture(L_S)}{\sum_{l \in L_S} |l|} \quad (4.34)$$

où $Couverture(L_S)$ correspond au nombre de documents distincts dans l'ensemble des listes présentées lors de la session S . Enfin, le nombre total de documents présentés dans cette même session est noté $|l|$.

- (b) Le ratio de couverture pertinente PCo analyse la pertinence moyenne des documents distincts retournés sur l'ensemble des sessions :

$$PCo = \frac{1}{|\theta|} \sum_{\tau \in \theta} \frac{1}{|\delta_\tau|} \sum_{S \in \delta_\tau} \frac{CouvPert(L_S)}{\sum_{l \in L_S} |l|} \quad (4.35)$$

où $CouvPert(L_S)$ correspond au nombre de documents distincts et pertinents dans l'ensemble des listes présentées lors de la session S .

4.3.1.3 Modèles de référence

Afin de comparer l'efficacité de nos contributions, nous définissons les trois modèles de référence –*baselines*– suivants :

- *SansDT* : ce modèle inclut les deux étapes de nos modèles, à savoir le calcul des scores basés sur les rôles et l'allocation des documents par la méthode d'apprentissage. Cependant, il ne garantit pas la division du travail qui vérifie que les listes de documents simultanément retournées aux collaborateurs sont distinctes.
- *SansEM* : ce modèle inclut seulement la première étape du calcul des scores basés sur les rôles tout en garantissant le paradigme de la division du travail. L'ordonnement collaboratif est alors déduit directement des scores basés sur les rôles, sans les optimiser.
- *SansEMDT* : ce modèle inclut la première étape de nos modèles, à savoir le calcul des scores basés sur les rôles, sans garantir la division du travail. Ce scénario peut être considéré comme un modèle de personnalisation dans un cadre de RI individuelle.

Dans le but de comparer notre contribution à un modèle de RIC de l'état de l'art, nous avons considéré le modèle de Foley and Smeaton (2009), noté FS :

- *FS* : ce modèle correspond à celui proposé par Foley and Smeaton (2009) et décrit dans la section 3.5.2.2. Afin de rendre le comparable à notre contribution, nous faisons l'hypothèse que l'indicateur d'autorité est similaire à la

notion d'expertise des collaborateurs. Deux versions, à savoir statique ou dynamique, de la mesure d'autorité sont définies. Nous détaillons ci-dessous la façon dont nous l'avons considéré dans le protocole d'évaluation de chacune de nos contributions.

1. *Modèle basé sur les rôles d'expert et de novice.* Compte tenu que notre modèle repose sur des rôles prédéfinis explicitement au début de la session, nous considérons la mesure d'autorité statique pour laquelle nous avons effectué un paramétrage optimal. Nous avons obtenu une valeur de 0,6 pour l'utilisateur le plus autoritaire, perçu dans notre cadre d'évaluation comme expert, et une valeur de 0,4 pour l'autre utilisateur, appelé ici novice.
2. *Modèle basé sur les rôles d'expert de sous-domaines.* A l'inverse du modèle de RIC basé sur les rôles d'expert et de novice de domaine où les niveaux d'expertise sont considérés verticalement et permettent d'assigner des rôles prédéfinis, nous faisons l'hypothèse que le facteur d'autorité estimé dynamiquement correspond plus à la présente configuration. En effet, la collaboration étant effectuée par un groupe d'experts, nous considérons que le facteur d'autorité dépend plus des jugements de pertinence effectués au cours de la session que de la différence des niveaux d'expertise, qui est plus thématique que hiérarchique.

4.3.2 Validation du modèle de RIC basé sur les rôles d'expert et de novice de domaine

Cette section présente les résultats obtenus pour l'analyse de l'efficacité de notre modèle d'ordonnancement basé sur les rôles d'expert et de novice de domaine. Notre évaluation expérimentale a pour objectif d'analyser l'efficacité classique de notre modèle en terme de précision, mais également en terme de diversité des résultats en lien avec la collaboration. A cet effet, nous orientons notre expérimentation selon les objectifs suivants :

1. Analyser l'efficacité globale de notre modèle au niveau de la session.
2. Analyser l'efficacité de notre modèle au niveau des rôles des collaborateurs afin d'étudier les différences éventuelles en terme d'efficacité de recherche entre les rôles d'expert et de novice de domaine.
3. Analyser l'impact des caractéristiques du groupe de collaborateurs sur l'efficacité de la session afin d'avoir une meilleure compréhension des différences des niveaux d'expertise sur l'efficacité de la session.

Nous présentons dans cette section les résultats obtenus selon une analyse de notre modèle basé sur les rôles d'expert et de novice de domaine au niveau de la session ainsi qu'au niveau des rôles. Pour chacune des analyses, nous présentons les résultats de notre modèle en fonction des deux méthodes de construction des paires de collaborateurs : exhaustive, notée *MExh*, et sélective, notée *MSel*.

4.3.2.1 Analyse de l'efficacité au niveau de la session de recherche collaborative

Dans cette section, nous analysons l'efficacité de notre modèle au niveau de la session de recherche à partir de 4 scénarios de référence qui tiennent compte des deux méthodes de construction des paires de rôles et des deux mesures d'estimation du niveau d'expertise des collaborateurs. Dans un premier temps, nous nous consacrons aux scénarios $S_{Exh-Spec}$ et $S_{Exh-Aut}$ pour lesquels les groupes de collaboration sont construits selon une méthode exhaustive et dont le niveau d'expertise est respectivement estimé selon un niveau de spécificité ou d'autorité. Nous dénombrons, pour chacun de ces scénarios, 243 paires de collaborateurs. Dans un second temps, nous analysons le cas particulier des groupes formés d'un expert et d'un novice de domaine, dont le niveau d'expertise est situé aux extrêmes du spectre de l'expertise. Ces scénarios sont générés par une méthode sélective et notés $S_{Sel-Spec}$ et $S_{Sel-Aut}$. Ces scénarios sont basés sur des mesures de spécificité et d'autorité afin d'estimer le niveau d'expertise des utilisateurs et regroupent respectivement 95 et 81 paires de collaborateurs.

Cas des groupes de collaboration basés sur une méthode exhaustive. Le Tableau 4.4 présente les précisions de recherche obtenues pour les différents modèles de RIC, selon les scénarios $S_{Exh-Spec}$ et $S_{Exh-Aut}$ générés par la stratégie de construction exhaustive des paires de collaborateurs et des deux indicateurs de niveau d'expertise basés sur la spécificité et l'autorité.

	Modèles	$P@30$	%Tx	$Co@30$	%Tx	$PCo@30$	%Tx
$S_{Exh-Spec}$	SansDT	0.260	+2.28%	0.383	+35.81% ***	0.086	+42.56% ***
	SansEM	0.250	+6.34% ***	0.364	+42.87% ***	0.080	+52.36% ***
	SansEMDT	0.283	-6.28%	0.279	+86.73% ***	0.058	+112.92% ***
	FS	0.221	+17.64% ***	0.442	+17.07% ***	0.086	+41.91% ***
	ENColl	0.266	-	0.520	-	0.123	-
$S_{Exh-Aut}$	SansDT	0.259	+1.86%	0.414	+37.88% ***	0.089	+45.17% ***
	SansEM	0.252	+4.98%*	0.396	+44.22% ***	0.082	+56.46% ***
	SansEMDT	0.284	-6.82%	0.273	+108.83% ***	0.053	+142.18% ***
	FS	0.216	+19.91% ***	0.442	+28.54% ***	0.086	+49.59% ***
	ENColl	0.265	-	0.571	-	0.129	-

TABLE 4.4: Comparaison de l'efficacité de notre modèle collaboratif au niveau des sessions pour les groupes créés de façon exhaustive - % Tx : Taux d'accroissement de notre modèle. Test de significativité de student * : $0.01 < t < 0.05$; ** : $0.001 < t < 0.01$; *** : $t < 0.001$

Les résultats montrent que les valeurs obtenues par notre modèle, noté *ENColl*, dépassent généralement celles des scénarios de référence, autant pour la mesure de précision, $P@30$, que pour les mesures orientées collaboration, $Co@30$ et $PCo@30$. Aucune différence notable n'est mise en évidence entre les deux

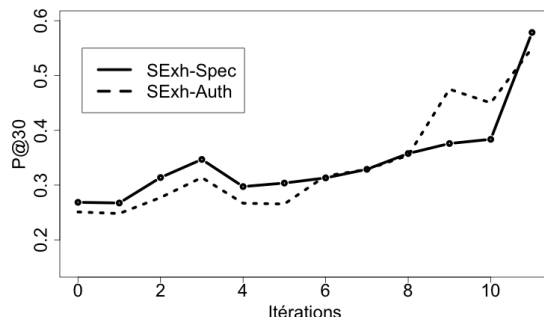


FIGURE 4.5: Analyse de la mesure de précision $P@30$ au niveau des itérations pour les groupes de collaboration construits à partir d’une méthode exhaustive

scénarios, ce qui souligne la robustesse de nos deux indicateurs dans l’identification des niveaux d’expertise. Plus précisément, nous constatons que notre modèle présente des taux d’accroissement significatifs par rapport au modèle collaboratif *FS* de 17.64% à 49.59%, quelles que soient les mesures d’évaluation ou les variantes des paires de rôles. Ces résultats peuvent être expliqués par le fait que notre modèle propose des listes personnalisées de documents en tenant compte du rôle des utilisateurs tandis que le modèle *FS* ordonne les documents selon un score agrégé de pertinence collective sans aucune personnalisation. Un autre aspect intéressant est d’analyser les différentes variantes de notre modèle, *SansDT*, *SansEM* et *SansEMDT*. Les résultats montrent que notre modèle fournit de meilleurs résultats que ceux des scénarios *SansDT* et *SansEM*. Cela suggère que (a) l’étape d’allocation des documents par le biais de l’algorithme EM est efficace dans l’apprentissage des scores des documents et (b) la division du travail permet d’améliorer la couverture des différents aspects du besoin en information partagé. Lorsque les deux composantes sont ôtées, nous constatons cependant une faible diminution non significative de notre modèle par rapport au scénario *SansEMDT*. Cependant, celle-ci est contrebalancée par le fait que les mesures orientées collaboration sont significativement plus élevées en faveur de notre modèle, avec des améliorations comprises entre 86.73% et 142.18%. Nous pouvons expliquer ce constat par le fait que le modèle *SansEMDT* propose un modèle d’ordonnancement personnalisé dans un contexte de RI classique tandis que notre modèle applique une méthode de personnalisation dans un contexte collaboratif, garantissant un double principe de division du travail, par l’algorithme d’allocation par méthode d’apprentissage et par le non-ordonnancement des documents simultanément présentés dans la liste de l’autre utilisateur. Ainsi la précision calculée pour notre modèle est plus particulièrement basée sur une pertinence résiduelle.

Nous poursuivons nos analyses en nous concentrant sur l’efficacité au niveau des itérations. Dans cet objectif, nous observons l’évolution de la mesure de micro-

précision au niveau de chaque itération k , décrite dans l'Equation 4.33. Les résultats sont illustrés dans la Figure 4.5 et montrent une tendance croissante des courbes avec une amélioration notable entre 115% et 119% du début jusqu'à la fin de la session, respectivement pour chacun des scénarios $S_{Exh-Spec}$ et $S_{Exh-Aut}$. Ce constat peut être expliqué par le fait que les scores des documents dépendent des deux rôles et des profils d'expertise basés sur les jugements de pertinence qui permettent d'affiner le besoin en information et, par conséquent, d'acquérir une meilleure perception de la pertinence des documents. Cependant, nous soulignons que l'amélioration à la onzième itération ne peut être considérée comme significative dans la mesure où le nombre de groupes atteignant cette itération est faible, à savoir seulement 4 groupes de collaboration sur les 243 pour les scénarios $S_{Exh-Spec}$ et $S_{Exh-Aut}$.

Cas des groupes de collaboration basés sur une méthode sélective. Nous considérons le cas particulier des groupes de collaboration caractérisés par une paire d'expert et de novice de domaine, explicitement définis par une méthode de construction sélective, dont le niveau d'expertise est estimé par un indicateur de spécificité $S_{Sel-Spec}$ ou d'autorité $S_{Sel-Aut}$. Le Tableau 4.5 présente les résultats obtenus pour notre modèle ainsi que pour les quatre modèles de référence.

	Modèles	$P@30$	%Tx	$Co@30$	%Tx	$PCo@30$	%Tx
$S_{Sel-Spec}$	SansDT	0.275	+4.09%*	0.362	+31.73%***	0.080	+29.63%***
	SansEM	0.268	+7.01%*	0.335	+42.46%***	0.072	+43.99%***
	SansEMDT	0.303	-5.26%	0.258	+84.73%***	0.050	+105.88%***
	FS	0,208	+32.21%***	0.429	+10.95%*	0.075	+37.99%***
	ENColl	0.287	-	0.477	-	0.103	-
$S_{Sel-Aut}$	SansDT	0.251	+0.86%	0.400	+36.44%***	0.081	+35.52%***
	SansEM	0.239	+5.87%	0.362	+50.11%***	0.070	+56.17%***
	SansEMDT	0.279	-9.29%	0.254	+114.48%***	0.048	+125.96%***
	FS	0,166	+51.20%***	0.429	+26.71%***	0.081	+34.22%***
	ENColl	0.253	-	0.544	-	0.110	-

TABLE 4.5: Comparaison de l'efficacité de notre modèle collaboratif au niveau des sessions pour les groupes créés de façon sélective - % Tx : Taux d'accroissement de notre modèle. Test de significativité de student * : $0.01 < t < 0.05$; ** : $0.001 < t < 0.01$; *** : $t < 0.001$

Les résultats montrent des tendances similaires à celles observées précédemment dans le Tableau 4.4. Plus particulièrement, nous notons qu'une amélioration globale en terme de précision est mise en avant pour notre modèle par rapport aux modèles de référence *SansDT*, *SansEM* et *FS*, ainsi qu'en terme de couverture avec une amélioration comprise 10.95% et 125.96% sur l'ensemble des modèles de référence. De plus, les résultats obtenus pour chacun de scénarios considérés confirment que l'estimation du niveau d'expertise, basé sur un critère de spécificité ou d'autorité, n'impacte pas l'efficacité de notre modèle. Nous observons

également que les tendances sont similaires entre les deux variantes de construction des paires de rôles. Notre modèle peut par conséquent être utilisé autant pour le cadre particulier de la collaboration entre un expert et un novice que pour le cas plus général de différence relative des niveaux d'expertise entre les deux collaborateurs.

Nous avons également analysé l'évolution de la micro-précision tout au long de la session au niveau des itérations. La Figure 4.6 illustre les résultats obtenus, mettant en évidence une croissance plus faible que celle présentée dans la Figure 4.5. Notons que la huitième itération est caractérisée par une chute de la micro-précision pour le scénario $S_{Sel-Spec}$ qui ne semble pas significative car un seul groupe parmi les 95 a effectué plus de 7 itérations.

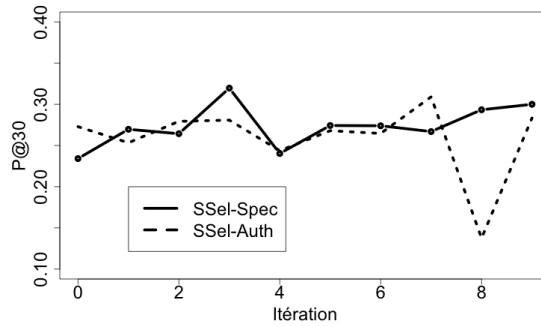


FIGURE 4.6: Analyse de la mesure de précision $P@30$ au niveau des itérations pour les groupes de collaboration construits à partir d'une méthode sélective

Les résultats obtenus dans cette section conduisent aux conclusions suivantes :

1. Notre modèle de RIC basé sur les rôles d'expert et de novice de domaine assure une meilleure diversité et pertinence des ordonnancements fournis aux collaborateurs tout au long de la session que les modèles de référence.
2. L'amélioration significative de notre modèle par rapport au scénario *SansEM* met en évidence l'importance de l'étape de l'allocation des documents basée sur l'algorithme EM. En effet, cette étape permet d'optimiser l'estimation des préférences des utilisateurs vis-à-vis des documents en tenant compte de leur niveau d'expertise.
3. Les scénarios *SansDT* et *SansEMDT* nous laissent suggérer que le paradigme de la division du travail peut conduire, a priori, à une faible réduction de l'efficacité d'un modèle de RIC. Cela peut être expliqué par le fait qu'un modèle collaboratif est plus orienté vers une précision résiduelle compte tenu de la configuration des sessions de recherche interactives et itératives. Cependant, les mesures de couverture et de couverture pertinente contrebalancent ce constat en mettant en avant le fait que notre modèle fournit plus de documents qui sont à la fois distincts et pertinents tout au long de la session tandis

que les modèles de référence sans division du travail favorisent la redondance des listes de documents. Par conséquent, nous pouvons supposer qu’une session de RIC optimale garantit à la fois pertinence et diversité des résultats, à l’inverse des scénarios individuels qui s’orientent particulièrement vers la satisfaction de la pertinence sans tout autant éviter la redondance dans les résultats retournés.

4. L’efficacité supérieure de notre modèle par rapport au modèle de l’état de l’art *FS* peut être expliquée par le fait que notre modèle intègre l’expertise des utilisateurs vis-à-vis du besoin en information dans l’objectif d’attribuer un document à l’utilisateur qui est le plus à même de le juger. Ainsi, nous garantissons que les documents retournés à un collaborateur correspondent au mieux à ses connaissances ; à l’inverse du modèle de Foley and Smeaton (2009) qui estime une pertinence collective basée sur l’ensemble des jugements de pertinence des collaborateurs.
5. Notre modèle est applicable dans différents contextes. La dualité de notre méthode de construction des groupes montre que notre modèle peut reposer sur des groupes avec des différences de niveaux d’expertise aussi bien relatives qu’absolues. En d’autres termes, il peut être appliqué dans le cas d’une paire de collaborateurs expert-novice mais aussi lorsqu’un utilisateur a plus d’expérience que son collaborateur sans nécessairement être expert du domaine. La méthode d’estimation du niveau d’expertise des collaborateurs utilisée dans notre expérimentation n’impacte pas la robustesse de notre modèle.

4.3.2.2 Analyse de l’efficacité au niveau des rôles des utilisateurs

Dans cette section, nous affinons notre analyse en focalisant nos expérimentations sur le facteur rôle des utilisateurs. A l’inverse de l’analyse précédente qui estime la pertinence globale de notre modèle sur l’ensemble de la session, cette section a pour objectif de mettre en valeur l’efficacité de notre modèle par rapport à chacun des rôles. Le Tableau 4.6 présente les mesures obtenues pour chacun des deux rôles en terme de micro-précision, $P@30^r$, estimée selon l’Equation 4.32, sous la condition de ne considérer que les listes de documents présentées à un rôle particulier.

Nous observons que notre modèle est généralement plus efficace, pour les deux rôles, en terme de micro-précision et de ratio de couverture pertinente par rapport aux autres scénarios de référence. Cependant, nous notons que seulement la moitié des améliorations est significative. Les conclusions que l’on peut déduire de ces résultats sont globalement les mêmes que celles obtenues dans la section précédente. D’une part, notre modèle fournit de meilleurs résultats, avec des améliorations significatives autour de 8%, que ceux obtenus pour le modèle de référence *SansEM*, renforçant nos déductions sur l’importance de l’étape de personnalisation proposée par l’algorithme d’apprentissage EM. D’autre part, le modèle de référence *SansEMDT* fournit de meilleurs résultats, généralement non significatifs, que ceux de notre modèle, pour les deux rôles utilisateurs. Comme

		$S_{Exh-Spec}$		$S_{Exh-Aut}$		$S_{Sel-Spec}$		$S_{Sel-Aut}$	
Modèles		$P@30^r$	%Tx	$P@30^r$	%Tx	$P@30^r$	%Tx	$P@30^r$	%Tx
Expert	SansDT	0.253	+5.43%	0.264	+5.67%	0.268	+9.12%*	0.285	+2.01%*
	SansEM	0.246	+8.44%*	0.259	+7.70%*	0.270	+8.22%	0.264	+9.78%
	SansEMDT	0.268	−0.59%	0.285	−2.30%	0.295	−0.94%	0.315	−7.87%
	FS	0.219	+22.11% **	0.233	+19.10%*	0.244	+19.74%*	0.234	+24.08%*
	ENColl	0.267		0.279		0.292		0.291	
Novice	SansDT	0.233	+0.60%	0.238	+1.67%	0.270	+0.47%	0.250	+4.11%
	SansEM	0.217	+7.70%*	0.227	+6.51%*	0.242	+11.68% ***	0.238	+8.97% ***
	SansEMDT	0.253	−7.30% ***	0.253	−4.52%*	0.287	−5.69%	0.262	−1.05.%
	FS	0.222	+5.63%	0.233	+3.86%	0.237	+14.17%	0.209	+23.81%
	ENColl	0.235		0.241		0.271		0.260	

TABLE 4.6: Comparaison de l'efficacité de notre modèle collaboratif au niveau des rôles - % Tx : Taux d'accroissement de notre modèle. Test de significativité de student * : $0,01 < t < 0,05$; ** : $0,001 < t < 0,01$; *** : $t < 0,001$

expliqué précédemment, l'application du paradigme de la division du travail, de par l'affichage de listes de documents distincts et l'allocation des documents basée sur l'algorithme EM, oriente notre modèle vers une précision résiduelle qui semble faire décroître la précision globale du modèle. Nous notons cependant que cette réduction est significative seulement pour les novices impliqués dans des groupes de collaboration construits par une méthode exhaustive. Cette observation est expliquée par le fait que la qualité de personnalisation des résultats dépend des rôles prédéfinis des collaborateurs plutôt que pour leur niveau d'expertise. En effet, étant donné que l'ensemble des paires de collaborateurs générées par la méthode sélective S_{Sel-*} est inclu dans l'ensemble des paires créées de façon exhaustive S_{Exh-*} , nous supposons que la différence en terme de significativité des résultats vient du défaut de personnalisation des ordonnancements vis-à-vis des collaborateurs qui ont des niveaux d'expertise proches. Cela met en valeur une explication qui justifie le fait que les jugements de pertinence issus à partir de listes de documents non personnalisées, particulièrement pour les novices, peut impacter négativement l'efficacité d'un modèle de recherche d'information. Enfin, les résultats de notre modèle sont plus élevés que ceux du modèle FS avec des améliorations significatives entre 19% et 24% pour l'expert de domaine et des améliorations, non significatives, comprises entre 3% et 23% pour le rôle de novice. Ce contraste peut être expliqué par les hypothèses sur lesquelles reposent le modèle FS . En effet, ce modèle d'ordonnement collaboratif ne considère pas de rôles utilisateurs et ne pratique aucune personnalisation des ordonnancements des documents. Même si la pertinence des documents est estimée au travers des jugements de pertinence des collaborateurs, le modèle applique un lissage dans la mesure où la pertinence est estimée de façon globale plutôt qu'au niveau individuel des collaborateurs. Par conséquent, le modèle FS traite les collaborateurs équitablement et compte tenu des résultats obtenus en comparaison de notre modèle, nous pouvons déduire que le modèle FS est

sensiblement aussi efficace que notre modèle l’est pour les novices.

D’une façon générale, nous notons que les accroissements obtenus grâce à notre modèle sont plus importants pour les experts que pour les novices. Ces résultats laissent supposer que le modèle proposé est plus efficace pour les experts que pour les novices avec une explication liée au fait que le jugement de pertinence (même implicite) de l’expert est une meilleure source d’évidence pour le processus de réinjection de pertinence que celui fourni par le novice, comme constaté également dans de précédents travaux (White and Dumais, 2009). Pour vérifier cela, nous effectuons une comparaison des moyennes de la mesure de micro-précision entre le rôle d’expert et celui de novice par l’intermédiaire d’un test de Student, illustrée dans le Tableau 4.7. Les résultats présentés dans ce tableau montrent clairement que les différences ne sont pas significatives (p-value > 0,05), ce qui ne nous permet pas de confirmer le biais constaté.

Scenarios	$S_{Exh-Spec}$	$S_{Exh-Aut}$	$S_{Sel-Spec}$	$S_{Sel-Aut}$
<i>Novice</i>	0.234	0.241	0.271	0.259
<i>Expert</i>	0.267	0.279	0.292	0.291
%Tx	+13.76%	+15.19%	+7.78%	+11.91%
p-value	0.20	0.16	0.58	0.38

TABLE 4.7: Comparaison des moyennes de l’efficacité de notre modèle par rapport aux rôles des collaborateurs - % Tx : Taux d’accroissement de notre modèle. Test de significativité de student * : $0.01 < t < 0.05$; ** : $0.001 < t < 0.01$; *** : $t < 0.001$

Enfin, nous examinons un aspect supplémentaire de notre modèle reposant sur l’hypothèse qu’un système optimal de RI permet aux utilisateurs non experts d’améliorer leurs connaissances du domaine (Hölscher and Strube, 2000; White and Dumais, 2009). Pour cela, nous assimilons l’expertise du domaine à la spécificité moyenne des documents inclus dans une liste en supposant qu’un novice de domaine améliore ses connaissances s’il est capable de lire des documents de plus en plus spécifiques. Dans l’objectif d’éviter le biais lié à la construction des groupes où les rôles sont attribués selon un niveau d’expertise basé sur l’indicateur de spécificité, nous effectuons cette analyse seulement sur les scénarios qui reposent sur une expertise estimée à partir de l’indicateur d’autorité. La figure 4.7 illustre l’évolution de la spécificité moyenne des listes retournées aux novices de domaine à chaque itération, représentée par les courbes $S_{Exh.Aut.Ret}$ et $S_{Sel.Aut.Ret}$, en comparaison de la spécificité des documents jugés comme pertinents, représentée par les courbes $S_{Exh.Aut.Sel}$ et $S_{Sel.Aut.Sel}$. L’itération 0 fait référence à l’étape d’initialisation S_{init} alors que les suivantes représentent les différentes itérations $k \in S_{iter}$ du modèle.

Dans un premier temps, nous nous consacrons à l’analyse de la spécificité moyenne des documents retournés à chaque itération afin d’évaluer dans quelle mesure notre modèle favorise l’amélioration de l’expertise des novices. A la fois

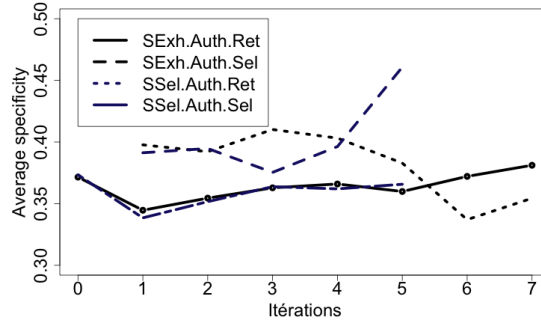


FIGURE 4.7: Analyse de l'évolution du niveau d'expertise des novices au cours de la session de recherche

pour les méthodes exhaustives et sélectives de construction des paires, nous observons une tendance croissante des deux courbes avec une baisse au niveau de l'itération 1. Celle-ci peut être expliquée par le fait que l'itération 1 correspond au début de la session à une absence de la personnalisation des listes de documents. En effet, les profils des deux utilisateurs sont identiques à l'itération 0 car ils n'intègrent que les termes de la requête. A l'itération 1, le profil est mis à jour et la personnalisation des listes est donc initiée. Ainsi, entre l'itération 1 et la dernière, l'accroissement de la spécificité moyenne est autour de 9% pour les deux listes. Cette tendance met en évidence le fait que notre modèle donne la possibilité aux novices d'améliorer leurs compétences de recherche, et particulièrement leurs connaissances vis-à-vis du besoin en information partagé, dans la mesure où notre modèle retourne des listes de documents dont le niveau de spécificité moyen augmente tout au long de la session.

Dans un second temps, nous souhaitons identifier si les connaissances des utilisateurs augmentent au fur et à mesure de la session. Pour cela, nous estimons la spécificité des jugements de pertinence effectués par le collaborateur novice à chaque itération. Nous observons que cette mesure décroît avec les itérations pour les groupes de collaboration qui ont été créés de façon exhaustive alors qu'elle augmente pour les groupes de collaboration générés par une méthode sélective. Ces observations soulignent le fait que les collaborateurs explicitement identifiés comme novice, avec un niveau d'expertise à l'extrême du spectre, ont plus de chance d'améliorer leur niveau d'expertise par rapport aux utilisateurs qui sont labélisés novices seulement parce qu'ils ont un niveau d'expertise inférieur à celui de leur collaborateur. En effet, les groupes formés par méthode exhaustive peuvent inclure deux experts avec des niveaux d'expertise sensiblement différents.

4.3.3 Validation du modèle de RIC basé sur un groupe d'experts de sous-domaines

Nous présentons et discutons dans cette section les résultats expérimentaux obtenus pour notre modèle d'ordonnancement collaboratif basé sur un groupe d'experts de sous-domaines, noté *ExpColl*. Etant donné que l'analyse des résultats du précédent modèle montre que les deux mesures d'expertise des utilisateurs sont assez robustes pour ne pas impacter l'efficacité du modèle proposé, nous considérons pour le modèle basé sur les experts de sous-domaines seulement la mesure du niveau d'expertise basée sur la spécificité moyenne des documents (Kim, 2006), présentée dans l'Equation 4.30.

Notre évaluation expérimentale a pour objectif d'analyser l'efficacité classique de notre modèle en terme de précision, mais également en terme de diversité des résultats en lien avec la collaboration. A cet effet, nous orientons notre expérimentation selon les objectifs suivants :

1. Analyser l'efficacité globale de notre modèle au niveau de la session.
2. Analyser l'impact de la taille des groupes sur l'efficacité globale de notre modèle.

Afin de déterminer les valeurs optimales des paramètres du modèle, nous adoptons une démarche de validation croisée selon deux ensembles. Pour cela, nous séparons les 277 sessions de recherche en deux sous-ensembles équivalents, notés Q_A et Q_B , comme illustré dans le Tableau 4.8. Dans ce qui suit, nous présentons successivement l'étape de paramétrage et les résultats obtenus en comparaison avec les modèles de référence.

	Sujets TREC	Nb. sessions
Q_A	303i, 326i, 339i, 352i, 357i, 362i, 408i, 428i, 446i	140
Q_B	307i, 322i, 347i, 353i, 365i, 366i, 387i, 392i, 414i, 431i	137

TABLE 4.8: Répartition des groupes d'experts pour la validation croisée

4.3.3.1 Paramétrage du modèle

Dans ce qui suit, nous expliquons la démarche en deux étapes effectuée pour attribuer une valeur aux paramètres du modèle qui maximise l'efficacité de notre modèle par rapport à l'ensemble des sujets TREC.

Diversification des documents. Dans un premier temps, nous avons fait varier le paramètre γ , présent dans l'Equation 4.18, qui a pour objectif de diversifier les résultats de recherche selon le score MMR (Carbonell and Goldstein, 1998). Le critère de diversité permettant de construire le sous-ensemble \mathcal{D}^* des $n = 1000$ documents est inversement proportionnel à la valeur de γ . Comme

illustré dans la Figure 4.8, nous notons que la valeur optimale qui maximise l'efficacité, estimée grâce à la précision moyenne ($MAP@30$) est égale à 1 pour les deux sous-ensembles Q_A and Q_B . Ce résultat est en cohérence avec l'analyse préliminaire pour valider que les sujets sont à multiples facettes.

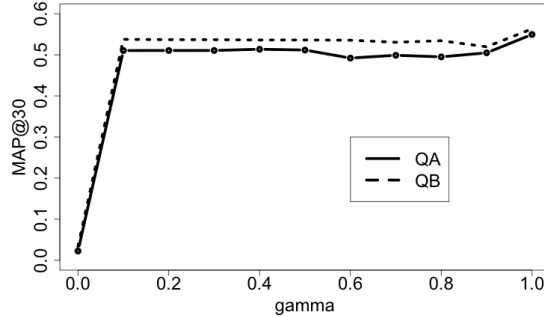


FIGURE 4.8: Paramétrage du facteur de diversité

Extraction des facettes de la requête. Dans un second temps, nous nous intéressons au nombre optimal de sujets permettant de représenter la distribution thématique des documents inclus dans l'ensemble diversifié, et par conséquent de la requête. Pour chacun des sujets TREC, nous avons appliqué l'algorithme du LDA sur les 1000 documents diversifiés en faisant varier le nombre de thèmes T de 20 à 200. Le nombre de thèmes est déterminé dans l'objectif de maximiser la vraisemblance, présentée dans l'Equation 4.19. Afin de rendre comparable les vraisemblances estimées, nous les avons normalisées pour chaque sujet TREC. La Figure 4.9 représente la valeur moyenne des vraisemblances normalisées en fonction du nombre de thèmes Ω extraits à partir de la collection pour chacun des deux ensembles de test. Comme illustré dans la Figure 4.9, nous retenons $\Omega = 200$ comme valeur optimale pour les deux sous-ensembles.

Dans leurs travaux, Deveaud et al. (2013) ont mis en évidence que le nombre de sous-thèmes d'une requête est généralement inférieur à 10. Ainsi, nous ajoutons un second critère, en plus de la vraisemblance, pour déterminer le nombre f de facettes parmi les 200 extraites. Pour cela, nous utilisons le score de pertinence d'un document présenté dans l'Equation 4.24, qui combine, grâce à un paramètre de pondération λ , un score basé sur la formule du BM25 et un score basé sur le LDA. La valeur de λ est déterminée de façon conjointe avec le nombre de facettes f les plus importantes. Nous avons ordonné les 1000 documents diversifiés en faisant varier les paramètres $\lambda \in [0; 1]$ et $f \in [1; 10]$ pour chacun des sujets TREC τ , comme illustré dans la Figure 4.10. Nous retenons les valeurs $\lambda = 0.6$ et $f = 5$ qui maximisent la mesure de précision moyenne ($MAP@30$) qui estime l'efficacité des ordonnancements pour les deux ensembles de requêtes Q_A et Q_B .

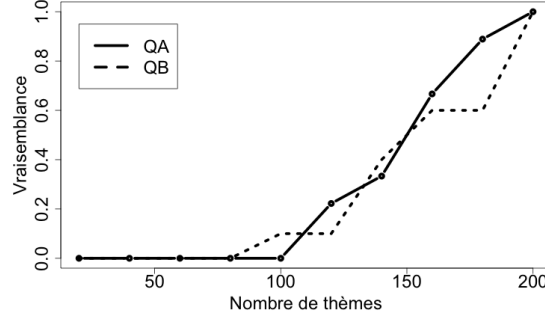


FIGURE 4.9: Paramétrage du nombre de thèmes issus de la collection à partir du LDA

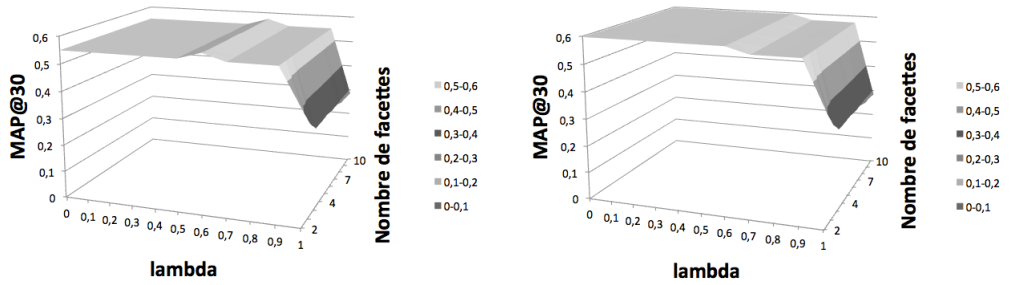


FIGURE 4.10: Paramétrage du nombre de facettes des requêtes collaboratives et de la fonction d'ordonnancement basée sur le LDA

4.3.3.2 Analyse de l'efficacité de la recherche

Dans cette section, nous présentons et discutons les résultats obtenus qui permettent de tester l'efficacité de notre modèle en comparaison aux modèles de référence. Le tableau 4.9 compare l'efficacité de notre modèle avec celle des trois modèles de référence détaillés dans la section 4.3.1.3.

Les mesures reportées montrent que notre modèle fournit généralement de meilleurs résultats. Plus particulièrement, la mesure de précision met en évidence que les résultats sont améliorés par notre modèle, en comparaison aux modèles qui ne considèrent pas l'étape d'ordonnancement collaboratif basé sur l'algorithme EM, à savoir *SansEM* et *SansEMDT*. Pour l'ensemble Q_B , nous notons des améliorations significatives par rapport au scénario *SansEM* d'une valeur de 20.60%. Ces résultats confirment l'importance de l'étape d'allocation des documents, comme nous avons pu le constater dans la section 4.3.2, par rapport à des modèles basés seulement sur la personnalisation, où les scores des do-

Apprentissage → Test		Modèles	<i>Co</i> @30	% <i>Ch</i>	<i>PCo</i> @30	% <i>Ch</i>	<i>P</i> @30	% <i>Ch</i>
$Q_B \rightarrow Q_A$	SansDT		0.486	+12.01	0.131	+19.58*	0.388	-0.30
	SansEM		0.571	-4.60	0.124	+26.73*	0.309	+20.60*
	SansEMDT		0.454	+17.56***	0.101	+55.40**	0.369	+1.13
	FS		0.545	+0.00	0.101	+55.40**	0.240	+35.48**
	ExpColl		0.545		0.157		0.372	
$Q_A \rightarrow Q_B$	SansDT		0.481	+10.78***	0.147	+4.29	0.414	-2.12***
	SansEM		0.526	+1.38	0.134	+14.77	0.379	+7.01
	SansEMDT		0.453	+17.69***	0.131	+17.16	0.394	+2.73
	FS		0.543	-1.88	0.129	+17.31*	0.298	+26.60*
	ExpColl		0.533		0.156		0.406	

TABLE 4.9: Comparaison de l'efficacité de notre modèle - % Tx : Taux d'accroissement de notre modèle. Test de significativité de student * : $0.01 < t < 0.05$; ** : $0.001 < t < 0.01$; *** : $t < 0.001$

cuments sont estimés selon le niveau d'expertise ou l'expertise de sous-domaines des utilisateurs. Comme constaté dans la section 4.3.2, nous observons que le scénario *SansDT* fournit de meilleurs résultats en terme de micro-précision, ce qui peut amener à conclure que la division du travail n'est pas un paradigme qui maximise l'effet synergique de la collaboration. Cependant, les mesures basées sur la notion de couverture contrebalancent ces résultats en fournissant des améliorations significatives atteignant 17.69% pour la métrique *Co*@30 et 55.40% pour celle de couverture pertinente *PCo*@30. Ces résultats confirment que notre modèle assure la diversité dans les ordonnancements des documents, estimée par la mesure *Co*@30, et la pertinence dans les ordonnancements diversifiés, estimée selon la mesure *PCo*@30. Le contraste entre la mesure de précision et celles basées sur la notion de couverture est expliqué par le fait que la première prend en compte la pertinence résiduelle dans les ordonnancements de documents tandis que la seconde ne considère pas la redondance entre les ordonnancements de documents.

En résumé et similairement aux conclusions mises en évidence dans la section 4.3.2.1, les résultats montrent que notre modèle collaboratif est plus approprié pour satisfaire les besoins en information à multiples facettes que des modèles de RI individuelle, modélisés à travers le scénario *SansEMDT*. En outre, les mesures basées sur la notion de couverture confirment que notre modèle favorise la diversité thématique dans les résultats de recherche grâce au principe de division du travail que nous avons mis en place par l'algorithme d'allocation des documents qui repose sur une classification améliorée par la condition de ne pas retourner un document simultanément à deux collaborateurs.

Nous poursuivons nos analyses par une étude de la micro-précision au niveau de l'itération, comme illustré dans la Figure 4.11. A l'inverse du modèle basé sur les rôles d'expert et de novice, nous observons ici que les sessions de RIC sont

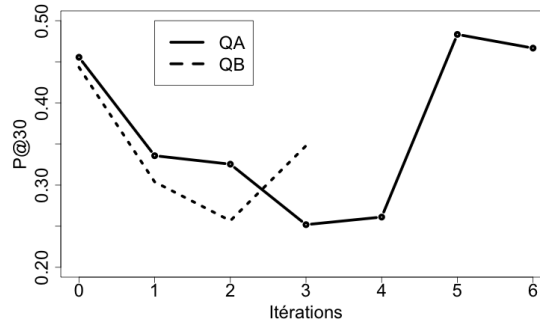


FIGURE 4.11: Analyse de la micro-précision au niveau des itérations

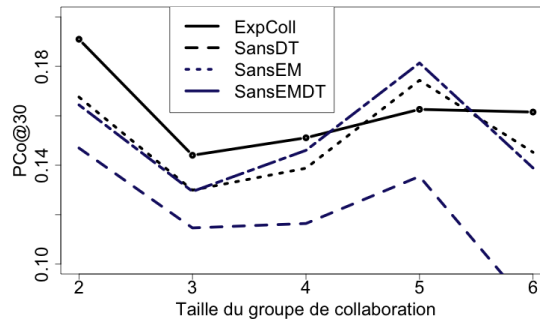


FIGURE 4.12: Impact de la taille des groupes de collaboration sur l'efficacité

moins longues. Par contre, similairement au modèle basé sur les rôles d'expert et de novice de domaine, nous constatons une chute de la précision entre l'étape d'initialisation et la première itération déclenchée pour personnaliser la liste des collaborateurs. La décroissance de l'efficacité semble durer plusieurs itérations, mais elle peut être expliquée par le fait que nous considérons des groupes de collaboration de plus de deux collaborateurs, pour la plupart, et que les premières itérations correspondent à la première itération de personnalisation de chacun des collaborateurs. Ainsi, en fin de session, la précision augmente, probablement du au fait que chacun des utilisateurs a pu constituer un profil qui permet à notre algorithme de garantir à la fois la résolution du besoin en information et la satisfaction des besoins individuels des collaborateurs.

Afin d'analyser plus en détail l'impact de la taille du groupe de collaboration sur l'efficacité de la session de recherche, nous représentons dans la Figure 4.12 la valeur moyenne de la mesure de couverture pertinente en fonction du nombre de collaborateurs au sein du groupe. Nous observons que la courbe représentant

notre modèle dépasse généralement celle des modèles de référence, à l'exception des groupes de taille 5. De plus, nous n'observons pas de réelle modification dans la tendance de la courbe pour des groupes plus large que la paire de collaborateurs, ce qui met en évidence la robustesse de notre modèle quelle que soit la taille des groupes de collaboration.

Résumé

Dans ce chapitre, nous avons proposé deux modèles d'ordonnancement de documents adaptés à une activité de recherche d'information collaborative où sont intégrés des rôles basés sur le niveau d'expertise des utilisateurs par rapport au sujet de la requête (Soulier et al., 2014b,c, 2013). Pour chacune de nos contributions, respectivement, nous avons distingué les niveaux d'expertise selon (a) une dimension verticale avec une paire de rôles d'expert et de novice de domaine et (b) une dimension horizontale avec un groupe d'experts de sous-domaines qui collaborent pour répondre à un besoin en information à multiples facettes.

Nos modèles incluent une première étape de calcul des scores des documents en fonction des rôles des utilisateurs et une seconde étape d'allocation des documents en utilisant une méthode d'apprentissage basée sur l'algorithme EM. La collaboration est supportée par une médiation système qui garantit la division du travail en distribuant les documents au collaborateur le plus à même de les juger et en évitant la redondance des actions entre collaborateurs.

L'évaluation expérimentale des deux modèles proposés est effectuée selon un cadre de simulation construit par extension d'un cadre d'évaluation de l'état de l'art et repose sur la collection TREC Interactive. Les résultats obtenus montrent l'intérêt de la personnalisation des scores de pertinence des documents selon l'expertise des collaborateurs, mais également l'apport de la méthode EM pour traduire la division du travail. La significativité des résultats a été attestée aussi bien sur la base des mesures de précision classique que des mesures orientées vers l'optimisation de la collaboration. De plus, la comparaison avec un modèle de l'état de l'art atteste le fait que notre modèle est particulièrement adapté à un contexte de recherche d'information collaborative, qui repose sur la complémentarité des compétences des utilisateurs (Sonnenwald et al., 2004; Spence et al., 2005).

Un des aspects intéressants de notre approche est qu'elle propose un cadre générique à la mise en place de modèles de RIC adaptables à d'autres paires de rôles. En effet, la première étape d'estimation des scores des documents selon les rôles intègre des facteurs liés aux rôles prédéfinis qui sont modulables selon les particularités des rôles souhaités. La deuxième étape peut être réutilisée afin d'optimiser les scores de documents pour la personnalisation des ordonnancements, comme défini dans notre contribution.

Discussion et conclusion

Dans cette partie, nous nous sommes intéressés à la prise en compte de l'expertise des utilisateurs dans un contexte de RIC. La principale motivation de nos deux contributions présentées repose sur la complémentarité des connaissances et compétences des utilisateurs (Sonnenwald et al., 2004; Spence et al., 2005) permettant ainsi de dégager l'effet synergique de la collaboration.

Dans un premier temps, nous avons considéré une distinction verticale des niveaux d'expertise au sein d'une paire de collaborateurs, conférant à ces derniers les rôles d'expert et de novice de domaine (Soulie et al., 2014b,c) (Section 4.2.2). Puis, nous nous sommes intéressés à une distinction horizontale entre un groupe d'experts de sous-domaines leur permettant de résoudre un besoin en information à multiples facettes (Soulie et al., 2013) (Section 4.2.3). Dans ces deux contributions, nous avons employé le même cadre de modélisation en deux étapes, permettant de prouver sa généralité dans la formalisation de modèles de RIC et ouvrant ainsi de nombreuses perspectives dans la mesure où il peut être adapté à d'autres rôles utilisateur. Notons cependant que nos modèles proposent d'ordonner les documents selon un niveau basé sur des itérations. Malheureusement, nous ne fournissons aucune liste de documents à la fin de la session qui synthétise les actions des collaborateurs et dégage la pertinence collective des deux collaborateurs vis-à-vis des documents. Par la définition du profil d'expertise dynamique qui évolue au fur et à mesure des jugements de pertinence des collaborateurs et traduit les compétences en terme de connaissance du domaine, nous avons également mis en évidence l'importance de la prise en compte dynamique des compétences et connaissances des utilisateurs. Également, cet aspect propose de nombreuses perspectives de recherche dans la mesure où le comportement des utilisateurs évolue au cours de la session, et par conséquent leurs compétences également, imposant ainsi aux modèles de RIC de s'adapter à cette dynamique.

Les évaluations expérimentales ont été réalisées selon un protocole de simulation de la collaboration basé sur la collection TREC Interactive. Nous avons considéré à la fois des mesures basées sur la pertinence et des métriques adaptées à la collaboration qui mesurent la couverture, appelées aussi diversité, des ordonnancements retournés au cours de la session. Les résultats sont prometteurs et permettent de montrer l'efficacité de nos contributions par rapport à ses différentes composantes et selon ses différents niveaux d'analyse.

Malgré ces nombreux atouts, nous sommes conscients des limites de nos contributions, que nous expliquons ci-dessous :

1. *Définition du profil d'expertise.* Nous avons proposé d'exprimer le niveau d'expertise des collaborateurs au travers d'un profil construit sur la base des termes des documents sélectionnés par les utilisateurs. Pour chacune de nos contributions, nous avons adopté une représentation différente : multinomiale pour les rôles d'expert et de novice de domaine et thématique pour les rôles d'expert de sous-domaines. Cette vision de l'expertise des utilisateurs peut être naïve dans la mesure où les termes sont considérés de façon indépendante alors qu'un traitement selon une approche connexionniste ou conceptuelle, similairement aux approches réalisées en RIP, permettraient de rajouter de la sémantique au profil d'expertise.
2. *Instanciation particulière de la RIC.* Nous avons défini un scénario de recherche basé sur un processus synchrone qui met à jour une liste de documents dès qu'un utilisateur juge un document comme pertinent. Nous soulignons que d'autres scénarios peuvent être considérés comme une session asynchrone où, particulièrement dans ce cas là, les interactions entre collaborateurs doivent être considérées dans le modèle d'ordonnancement afin d'analyser le flux d'information entre les collaborateurs. Notons que la prise en compte des interactions peut également être un facteur important car les interactions mettent en évidence les stratégies de recherche et/ou affinent le besoin en information partagé des collaborateurs. Egalement, les ordonnancements des documents peuvent être mis à jour seulement si l'utilisateur reformule une nouvelle requête ou considère qu'elle ne contient plus d'éléments pertinents.
3. *Limite des expérimentations basées sur la simulation.* Nos protocoles d'évaluation sont basés sur la simulation de la collaboration à partir de fichiers log de sessions interactives, offrant l'avantage de mise en place d'expérimentations de moindre coût mais qui, en contrepartie, limitent la portée des résultats qui en sont issus. En recherche d'information individuelle, cette pratique a été adoptée dans de nombreux travaux (Keskustalo et al., 2006; White et al., 2005). Nous sommes conscients cependant que la simulation comporte certaines faiblesses car des fichiers log de sessions individuelles ne reflètent pas réellement le comportement des utilisateurs en contexte de recherche collaborative. Par exemple, la simulation de la session ne peut comporter aucune trace des interactions, des communications ou stratégies de recherche que les collaborateurs auraient pu éventuellement mener.

Dans la partie suivante, nous nous intéressons à l'aspect de la dynamique des comportements évoqués précédemment selon l'axe de modélisation de ces comportements par les rôles. Plus particulièrement, nous souhaitons poser certains fondements quant à la considération des rôles dans une session de RIC vis-à-vis des comportements des utilisateurs afin d'assurer une complémentarité optimale de leurs compétences.

Troisième partie

Vers des modèles de recherche d'information collaborative basés sur une médiation hybride utilisateur-système

Introduction

Golovchinsky et al. (2009b) proposent une taxonomie des rôles de collaboration basée principalement sur la notion d’expertise qui peut être considérée selon une approche thématique, comme nous l’avons abordée également dans le chapitre précédent, ou une approche plus technique qui recouvre les compétences de recherche des collaborateurs. Dans cette partie, nous nous intéressons particulièrement à cet aspect en faisant l’hypothèse qu’une session de RIC peut être optimisée si elle exploite la complémentarité des compétences de recherche des collaborateurs. L’intuition qui guide notre approche repose principalement sur une des motivations majeures de la collaboration qui s’appuie sur l’insuffisance des compétences d’un utilisateur unique, et par conséquent, du besoin de rassembler des compétences complémentaires pour résoudre un besoin en information partagé (Sonnenwald et al., 2004; Spence et al., 2005). De plus, la complémentarité des compétences peut permettre une division du travail aisée dans la mesure où chacun des collaborateurs peut structurer ses actions (Kelly and Payne, 2013) et se spécialiser vers des actions spécifiques aux compétences pour lesquelles il est le plus efficace.

De nombreux travaux (Pickens et al., 2008; Shah et al., 2010) se sont intéressés à la structuration de la collaboration à travers des rôles distingués par des tâches différentes. Cependant, nous suggérons de nombreuses limites sous-jacentes à ces contributions :

1. Les rôles prédéfinis, et plus particulièrement les tâches associées, assignés aux collaborateurs peuvent ne pas correspondre aux compétences intrinsèques des utilisateurs, et par conséquent, biaiser leur comportement de recherche.
2. Les rôles sont figés sur l’ensemble de la session et ne prennent pas en compte le fait qu’un collaborateur peut avoir plusieurs rôles suivant ses interactions avec les collaborateurs et les sources d’information (Imazu et al., 2011). Par exemple, il peut, dans un premier temps, recourir à un rôle de “*requêteur*” et utiliser ces compétences liées à la formulation de requêtes afin de récolter le maximum d’informations, et, par la suite, devenir “*analyste de l’information*” en favorisant un processus de *sensemaking*.

La limite liée à la dynamique des rôles peut être partiellement résolue dans un contexte de la RI individuelle où certains auteurs proposent des techniques d’analyse du comportement des utilisateurs par le biais de leurs jugements de pertinence (Agichtein et al., 2006; Joachims, 2002; Radinsky et al., 2013) ou d’apprentissage des rôles latents des utilisateurs dans des réseaux sociaux ou de communication, sur la base de leurs interactions (Henderson et al., 2012; McCallum et al., 2007). Cependant, ces modèles ne permettent pas d’optimiser l’efficacité de la session en proposant d’adapter les ordonnancements à la fois par rapport aux compétences complémentaires des utilisateurs et aux paradigmes de la collaboration.

Dans ce contexte, nous nous intéressons, particulièrement dans cette partie du manuscrit, à la problématique d’évolution et de complémentarité du comporte-

ment des collaborateurs qui construisent leurs rôles implicites au cours d'une session de RIC. Ainsi, nous proposons des modèles dynamiques d'apprentissage des rôles des collaborateurs basés sur l'hypothèse que les collaborateurs sont capables de structurer la collaboration par le biais de rôles implicites afin de bénéficier, tout au long de la session, des compétences pour lesquelles ils sont les plus efficaces. En plus de bénéficier de la complémentarité des connaissances des collaborateurs, notre principale motivation est de garantir le principe du bénéfice mutuel en utilisant un modèle d'ordonnancement adapté aux rôles des collaborateurs. La session de recherche est alors perçue comme un élément hybride qui repose à la fois sur (a) une médiation utilisateur, laissant les collaborateurs libres de leurs actions et (b) une médiation système, adaptant les ordonnancements aux rôles des collaborateurs détectés à partir de leurs actions.

Cette partie s'articule autour de deux contributions :

1. Un modèle d'apprentissage supervisé et dynamique des rôles des collaborateurs qui considère un ensemble prédéfini de rôles candidats pouvant être assignés aux collaborateurs tout au long de la session (Chapitre 5). Ce chapitre inclut au préalable une analyse préliminaire qui met en évidence les hypothèses de recherche qui ont motivé notre contribution.
2. Un modèle d'apprentissage non-supervisé et dynamique des méta-rôles des collaborateurs qui se base sur l'hypothèse que les compétences des collaborateurs ne sont pas toujours en parfaite adéquation avec les caractéristiques des rôles prédéfinis (Chapitre 6). Ce modèle est une généralisation du modèle présenté dans le Chapitre 5.

Chapitre 5

Modèle d'apprentissage supervisé des rôles en contexte de collaboration

Introduction

La complémentarité des compétences des utilisateurs est une des principales motivations à la collaboration lors de la résolution d'une tâche complexe (Sonnenwald et al., 2004; Spence et al., 2005). Un des moyens d'exploiter les compétences des collaborateurs repose sur la définition de rôles adaptés qui, de plus, permettent de structurer la collaboration (Kelly and Payne, 2013). Dans ce chapitre, nous nous intéressons à la problématique d'assignation de rôles de collaboration prédéfinis, comme ceux présentés dans la taxonomie de Golovchinsky et al. (2009b), à une paire de collaborateurs afin de bénéficier des compétences de recherche pour lesquelles ils sont les plus efficaces. Ainsi, nous proposons un modèle d'apprentissage supervisé des rôles des collaborateurs reposant sur une session de recherche perçue comme un élément hybride entre une médiation système et une médiation utilisateur (Soulé et al., 2014a). L'objectif principal de notre approche est d'optimiser l'efficacité de la session de RIC en adaptant les modèles d'ordonnancement aux rôles des collaborateurs.

La suite du chapitre est organisée comme suit. La section 5.1 présente les motivations de notre contribution et introduit les questions de recherche sous-jacentes. La section 5.2 définit les notations et définitions utilisées par la suite. Notre modèle d'apprentissage des rôles est présenté dans la section 5.3. Enfin, l'évaluation expérimentale menée dans l'objectif de valider notre modèle est détaillée dans la section 5.4.

5.1 Motivations et questions de recherche

5.1.1 Contexte

Les approches de médiation système permettent d’ordonner collaborativement les documents afin de résoudre un besoin en information partagé. Ces approches sont particulièrement efficaces lorsqu’elles permettent de garantir le principe du bénéfice mutuel nécessaire à l’implication des collaborateurs dans la session dans la mesure où les ordonnancements de documents permettent à la fois de répondre à leurs attentes et au besoin en information.

D’une part, certaines approches en RIC fournissent une solution basée sur les rôles symétriques de pairs (Foley and Smeaton, 2009; Morris et al., 2008). Le bénéfice mutuel est alors assuré dans la mesure où les utilisateurs sont, par conséquent, totalement libres de structurer leurs actions. Cependant, une des principales limites de ces approches est qu’elles n’adaptent pas les ordonnancements aux compétences des collaborateurs et se limitent à un appariement basé sur le contenu des documents, qu’il soit personnalisé (Morris et al., 2008) ou non (Foley and Smeaton, 2009).

D’autre part, une autre technique (Pickens et al., 2008; Shah et al., 2010) consiste en l’assignation de rôles fixes et prédéfinis afin de diviser le travail entre les utilisateurs (Kelly and Payne, 2013). Les collaborateurs sont alors guidés dans le processus de recherche et les fonctions d’ordonnement sont améliorées par la mise en place de techniques de RI, telles que la reformulation de requêtes, dans la mesure où les ordonnancements de documents respectent les tâches de recherche sous-jacentes aux rôles. La principale limite de ces approches réside dans la stabilité des assignations de rôles inhibant la possibilité à un collaborateur d’utiliser plusieurs compétences, en d’autres termes d’être assimilé à plusieurs rôles tout au long de la session. En effet, Hendaheva and Shah (2013) se sont intéressés aux stratégies de recherche employées par les utilisateurs dans le cadre d’une session de RI individuelle afin de répondre à un besoin en information et ont mis en évidence leur évolutivité sur l’ensemble de la session.

A partir de cette brève analyse des travaux de l’état de l’art, nous pouvons formuler une première intuition qui repose sur le fait qu’une session de RIC est optimale lorsqu’elle permet de résoudre le besoin en information en considérant les compétences complémentaires des collaborateurs. L’objectif est ainsi d’attribuer dynamiquement aux collaborateurs le meilleur rôle qui les caractérise, permettant ainsi de structurer la collaboration en temps réel (Kelly and Payne, 2013). Afin de valider notre intuition, nous présentons dans ce qui suit une analyse préliminaire effectuée sur un jeu de données basé sur des sessions de collaborations réelles. Puis, nous introduisons les notations et notre modèle d’apprentissage des rôles dans un contexte collaboratif. Enfin, nous présentons le protocole expérimental mené pour évaluer l’impact de notre modèle d’apprentissage des rôles sur l’efficacité d’une session de RIC.

5.1.2 Analyse préliminaire

Afin de valider notre intuition sur l'exploitation de la complémentarité et de la dynamique des compétences des collaborateurs, nous avons effectué une analyse préliminaire dont l'objectif est d'étudier le comportement des collaborateurs lors d'une session de recherche. Plus particulièrement, nous nous intéressons à la façon dont les utilisateurs structurent la collaboration et mettent en avant la complémentarité de leurs compétences selon trois scénarios :

1. GS : une session de recherche où les utilisateurs ont été assignés aux rôles de Rassembleur ou de Surveilleur –*Gatherer-Surveyor*– (Shah et al., 2010).
2. PM : une session de recherche où les utilisateurs ont été assignés aux rôles de Prospecteur ou de Mineur –*Prospector-Miner*– (Pickens et al., 2008).
3. SansRole : une session de recherche sans contrainte laissant libres les collaborateurs de s'organiser au sein du processus de RIC.

Cette analyse repose sur une expérimentation utilisateurs que nous décrivons dans ce qui suit.

5.1.2.1 Méthodologie expérimentale

Nous avons effectué une expérimentation utilisateurs regroupant 75 paires de collaborateurs qui ont travaillé sur une tâche de recherche exploratoire. Nous décrivons dans ce qui suit les participants, la tâche de recherche, le besoin en information ainsi que le déroulement de l'expérimentation.

Les participants. Les participants ont été recrutés à l'Université Paul Sabatier et sont essentiellement des étudiants de niveau licence, master ou doctorat, âgés entre 18 et 30 ans. Lors du recrutement, nous nous sommes assurés que les participants avaient une expérience dans la navigation web et dans l'utilisation des moteurs de recherche. De plus, nous avons recruté les participants par paires en s'assurant qu'ils se connaissaient au préalable, afin de faciliter la collaboration entre les membres du groupe. Chaque groupe de participants a été assigné à un scénario de recherche.

La tâche de recherche. La tâche de recherche consistait en une recherche exploratoire où les participants naviguaient sur le web et soumettaient des requêtes aux moteurs de recherche publics, principalement Google¹. L'objectif des participants était d'identifier le maximum de pages web pertinentes en rapport avec un sujet dans une session de 30 minutes dans l'idée d'écrire un rapport ensemble à partir des informations sauvegardées durant la tâche de recherche collaborative. Afin de sauvegarder les informations identifiées pendant la session de RIC, les participants avaient à disposition un système de collaboration implémenté

1. <http://www.google.com>

sous la forme d'un plugin Firefox. Ce plugin est construit sur la base du système de collaboration Coagmento (Shah and González-Ibáñez, 2011b) dont une partie du code est disponible en open-source² et que nous avons adapté à nos besoins. Cet outil est composé de deux éléments :

1. Une barre d'outil (Figure 5.1) permettant aux collaborateurs de noter, d'annoter des pages web ou de sauvegarder des morceaux d'information, appelés également *snippets*.



FIGURE 5.1: Barre d'outils du système de collaboration

2. Une barre latérale (Figure 5.2) composée d'un système de messagerie instantanée et d'un espace de visualisation des documents sélectionnés par les participants.



FIGURE 5.2: Barre latérale du système de collaboration

En outre de fournir un support à la collaboration entre les participants, ce plugin nous a permis d'enregistrer les actions des utilisateurs ainsi que leur méta-données, à savoir les pages visitées/annotées/notées, les snippets, les requêtes soumises et les messages échangés entre participants.

Le besoin en information. Le sujet de chacun des scénarios a été extrait de la collection TREC Interactive 6-7-8, dont la tâche des participants à cette campagne d'évaluation consiste en l'identification des différents aspects du sujet (Over, 2001). Cette collection est également utilisée dans la littérature pour l'évaluation de modèles de RIC (Foley and Smeaton, 2009; Soulier et al., 2014b,c,

2. <https://github.com/InfoSeeking/CoagmentoCollaboratory>

2013). Afin de limiter le biais des connaissances des utilisateurs et de s’assurer que les participants étaient sensibilisés au sujet, nous avons identifié trois sujets associés à chacun des scénarios en s’assurant de l’homogénéité de la difficulté de la tâche exploratoire, comme illustré dans la Figure 5.3. Pour cela, nous avons estimé la mesure de rappel basée sur la notion d’aspects –*aspectual recall*– du besoin en information (Over, 2001) (et non de pages pertinentes) sur l’ensemble des sujets de la collection TREC Interactive. Les sujets homogènes, assignés aux participants avec leur accord, sont caractérisés par des mesures d’*aspectual recall* proches, qui, de plus, gravitent autour de la médiane.

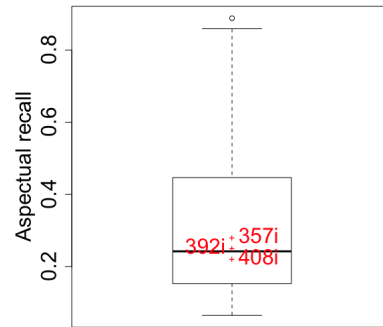


FIGURE 5.3: Difficulté de la tâche exploratoire

Nous avons ainsi retenu les sujets suivants :

1. “Extinction de la faune” (sujet 347i “*Wildlife extinction*” TREC-6 Interactive) pour le scénario GS.
2. “Robotique” (sujet 392i “*Robotic*” TREC-7 Interactive) pour le scénario PM.
3. “Tempêtes tropicales” (sujet 408i “*Tropical storm*” TREC-8 Interactive) pour le scénario SansRole.

Chacun des sujets a été reformulé en y ajoutant un contexte, l’objectif de la tâche, le sujet TREC et la tâche collaborative, comme illustré ci-dessous :

“<Contexte de la tâche> Le maire de votre village de campagne doit décider s’il construit un immense complexe industriel ou s’il développe un parc classé réserve naturelle. Vous êtes tous les deux garde forestier et vous souhaitez soulever le problème des animaux en voie d’extinction, généré par le développement abusif des villes et des comportements non respectueux des hommes.. </Contexte de la tâche> <Aspect collaboratif> Afin de construire un discours bien argumenté sur le sujet, vous souhaitez acquérir plus de connaissance et collecter différentes informations dans l’objectif de produire ensemble un rapport. </Aspect collaboratif> <sujet TREC> Vous traiterez donc le sujet des « animaux en voie d’extinction » et, entre autres : des principales espèces concernées, des efforts faits par les pays pour les préserver, dans quelles mesures les hommes sont-ils concernés par ce problème, ... </sujet TREC> <Tâche collaborative> Durant

Categ.	Description	Exemples
Page	Toutes les déclarations concernant la pertinence des pages	Identifier l'information pertinente dans des pages contenant beaucoup d'information
Comm.	Toutes les déclarations à propos de soucis de communication	Echange par le chat moins pratique qu'en face-à-face
Org.	Toutes les déclarations à propos de l'organisation de la tâche	Diviser les tâches de recherche
Topic	Toutes les déclarations à propos du sujet de la tâche	Se focaliser sur la tâche et ne pas regarder d'autres sujets
Technic.	Toutes les déclarations reliées à des problèmes techniques	La fenêtre du chat est trop petite
Role	Toutes les déclarations reliées à l'attribution des rôles	Ce que j'ai trouvé difficile était de suivre le rôle

TABLE 5.1: Catégorisation des réponses des participants à propos de la nature de la difficulté de la tâche

la session, vous devez identifier tous les documents ou morceaux de documents qui vous semblent traiter le sujet en sauvegardant les documents, les snippets (parties du texte), et annotant les documents. Merci de prendre en considération que cette session est une tâche préliminaire à l'écriture d'un rapport et que toutes les informations que vous collectez doivent aider à construire ce rapport.</Tâche collaborative>

Déroulement de l'expérimentation. Nous avons mis en place une expérimentation qui repose sur quatre étapes principales :

1. Une étape de saisie d'un formulaire de recrutement qui a permis d'enregistrer les participants à l'expérimentation et de collecter des informations démographiques, telles que leur âge, leur sexe ou leur habitude de navigation sur le web. Pour les participants impliqués dans des scénarios basés sur les rôles, à savoir PM et GS, des questions ont été respectivement rajoutées dans l'objectif d'attribuer le rôles le plus ressemblant aux caractéristiques des participants :
 - (a) Pour le scénario GS, nous avons posé la question suivante : "Quel participant au sein de votre groupe de collaboration correspond le plus à chacun de ces deux comportements : (i) regarder rapidement les pages web et (ii) naviguer vers des documents pas forcément en rapport avec le coeur du sujet ?".
 - (b) Pour le scénario PM, nous avons formulé la question suivante : "Quel participant au sein de votre groupe de collaboration correspond le plus à chacun de ces deux comportements : (i) identifier les documents les plus pertinents et (ii) naviguer vers des documents pas forcément en rapport avec le coeur du sujet ?"

	PM	GS	SansRole
Sujet	Robotique	Extinction de la faune	Tempête tropicale
Nombre de paires	24	26	25
Nombre total de pages web visitées	5137	4884	4734
Nombre total de pages web sauvegardées/notées	1195	941	333
Nombre total de snippets sauvegardés	340	588	306
Nombre total de requêtes soumises	1686	1633	1174
Nombre moyen de mots par requête	3.59	2.89	3.65

TABLE 5.2: Statistique des expérimentations utilisateurs

2. Une étape de sensibilisation sur la base d'un tutoriel dans lequel nous avons présenté les composantes du système et où les participants ont pu tester le système pendant quelques minutes. De plus, dans le cas d'un scénario basé sur des rôles prédéfinis, nous avons donné dans un premier temps, sous format papier, les spécificités de chacun des rôles assignés aux participants, que nous avons ensuite expliquées oralement afin de s'assurer de la compréhension des participants vis-à-vis des contraintes des rôles.
3. Une étape de réalisation de la tâches de RIC au travers d'une session qui a duré 30 minutes dont l'objectif était de réaliser la tâche exploratoire.
4. Une étape de saisie de questionnaire de retour où nous avons pu collecter les commentaires et les impressions des participants suite à la réalisation de la tâche. Parmi les questions posées, nous nous sommes intéressés au niveau de difficulté de la tâche (facile, moyen, difficile) et au type de difficulté. Pour ce dernier aspect, la réponse était libre, mais nous avons procédé a posteriori à une catégorisation manuelle par deux assessseurs (un étudiant de master et un doctorant) afin d'assurer la cohérence des réponses. La catégorisation utilisée est présentée dans le Tableau 5.1.

Pour un meilleur aperçu des caractéristiques des expérimentations, le Tableau 5.2 présente quelques statistiques sur chacun des jeux de données.

5.1.2.2 Notations : modélisation de la session collaborative

Dans cette section, nous présentons les notations relatives à la modélisation de la session de recherche d'un utilisateur, nécessaire à notre analyse.

La session de recherche $S_{u_j}^{(t_l)}$ d'un utilisateur u_j regroupe l'ensemble des requêtes $q_{u_j}^{(1)}, q_{u_j}^{(2)}, \dots, q_{u_j}^{(t_l-1)}$, reformulées jusqu'à l'instant t_l ($0 < t_l < T$) et ordonnées de façon chronologique, ainsi que les ensembles des résultats associés $D_{u_j}^{(1)}, D_{u_j}^{(2)}, \dots, D_{u_j}^{(t_l-1)}$. Etant donné que le facteur le plus important pour détecter le rôle des collaborateurs repose sur la complémentarité des compétences entre

		Description	Mesure
C_R	N_q	Nombre de requêtes	Nombre de requêtes soumises
	L_q	Longueur des requêtes	Nombre moyen de mots par requête
	R_q	Recouvrement entre requêtes	Ratio moyen de mots communs entre deux requêtes successives
C_P	N_{pq}	Nombre de pages par requête	Nombre moyen de pages visitées par requêtes
	D_p	Durée de visite	Durée moyenne en secondes d'une visite sur une page
C_J	N_j	Nombre de jugements de pertinence	Nombre de jugements de pertinence (annotations, notations et snippets)
C_M	N_m	Nombre de messages	Nombre de messages échangés entre les participants du groupe

TABLE 5.3: Caractéristiques de comportement utilisées pour l'analyse préliminaire des comportements

collaborateurs, nous favorisons une modélisation de la session selon une représentation basée sur des caractéristiques de comportement $F = \{f_1, \dots, f_k, \dots, f_n\}$ capturées durant la session de recherche, similairement à Agichtein et al. (2006) et White and Dumais (2009). Elles sont décrites dans le Tableau 5.3. Nous les avons regroupés en quatre catégories C_R , C_P , C_J et C_M qui recensent respectivement des caractéristiques liées aux requêtes, aux pages, aux jugements de pertinence et aux messages échangés entre participants. Chacune d'elles permet d'inférer le comportement et les stratégies des utilisateurs durant la session.

Dans ce contexte, la session $S_{u_j}^{(t_i)}$ d'un utilisateur u_j est représentée par une matrice $S_{u_j}^{(t_i)} \in \mathcal{R}^{t_i \times n}$ où chaque élément $S_{u_j}^{(t_i)}(t_{l'}, f_k)$ représente la valeur cumulée de la caractéristique de comportement f_k pour l'utilisateur u_j agrégée durant la session de recherche durant l'intervalle temporel $[0 \dots t_{l'}]$. Nous avons fait le choix d'agréger les valeurs des caractéristiques pour un utilisateur u_j à un instant $t_{l'}$ afin de mettre en évidence son comportement global au détriment de l'analyse d'un comportement ponctuel des collaborateurs où il peut exister un biais par la prise en compte des actions bruitées. De plus, elles permettent de considérer le comportement évolutif des utilisateurs.

Afin de comparer le comportement entre deux collaborateurs u_1 et u_2 par rapport à leur session de recherche, nous définissons également la matrice de différences $\Delta_{u_1, u_2}^{(t_i)} \in \mathcal{R}^{t_i \times F}$. Chaque élément $\Delta_{u_1, u_2}^{(t_i)}(t_{l'}, f_k)$ est défini par la différence de la valeur de l'utilisateur u_1 et de l'utilisateur u_2 vis-à-vis de la caractéristique de comportement f_k à l'instant $t_{l'}$:

$$\Delta_{u_1, u_2}^{(t_i)}(t_{l'}, f_k) = S_{u_1}^{(t_i)}(t_{l'}, f_k) - S_{u_2}^{(t_i)}(t_{l'}, f_k) \quad (5.1)$$

Ces notions sont abordées dans l'Illustration 5.1.

Illustration 5.1 Modélisation de la session de recherche d'un utilisateur.

Soit une session collaborative de 30 minutes impliquant deux utilisateurs u_1 et u_2 cherchant à identifier des informations sur la thématique de “*global warming*”. Nous considérons que les collaborateurs ont réalisé les actions suivantes pendant les cinq premières minutes de la session :

u	t	actions	compléments d'information
u_2	0	requête soumise	“global warming”
u_1	1	requête soumise	“global warming”
u_2	8	document d_1 visité	commentaire “interesting”
u_2	12	document d_2 visité	
u_2	17	document d_3 visité	noté 4/5
u_2	19	document d_4 visité	
u_1	30	requête soumise	“greenhouse effect”
u_1	60	requête soumise	“global warming definition”
u_1	63	document d_{20} visité	noté 3/5
u_1	70	requête soumise	“global warming protection”
u_1	75	document d_{21} visité	
u_2	100	document d_5 visité	noté 5/5
u_2	110	document d_6 visité	noté 4/5
u_2	120	document d_7 visité	
u_1	130	requête soumise	“gas emission”
u_1	132	document d_{22} visité	noté 4/5
u_2	150	document d_8 visité	
u_2	160	document d_9 visité	
u_2	170	document d_{10} visité	
u_2	200	document d_{11} visité	commenté “great”
u_2	220	document d_{12} visité	
u_2	240	document d_{13} visité	
u_1	245	requête soumise	“global warming world protection”
u_1	250	requête soumise	“causes temperature changes”
u_1	298	requête soumise	“global warming world politics”

La matrice $S_{u_j}^{(t_l)}$ de comportement des deux utilisateurs $u_j \in \{u_1, u_2\}$ à l'instant $t = 300$, soit 5 minutes, où une ligne correspond à un instant $t_l \leq t_l$ et une colonne à une caractéristique de comportement $f_k \in F$ est extraite comme suit. Nous choisissons de définir des intervalles de temps réguliers de 60 secondes pour identifier le comportement des utilisateurs. Ce dernier est modélisé par un ensemble de quatre caractéristiques $\{N_q, N_d, N_c, N_n\}$ correspondant respectivement au nombre de requêtes, de documents, de commentaires et de notes. La modélisation de chacun des utilisateurs u_1 et u_2 respectivement est la suivante :

$$S_{u_1}^{(300)} = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 4 & 2 & 0 & 1 \\ 5 & 3 & 0 & 2 \\ 5 & 3 & 0 & 2 \\ 8 & 3 & 0 & 2 \end{pmatrix} \text{ et } S_{u_2}^{(300)} = \begin{pmatrix} 1 & 4 & 1 & 1 \\ 1 & 7 & 1 & 3 \\ 1 & 10 & 1 & 3 \\ 1 & 13 & 2 & 3 \\ 1 & 13 & 2 & 3 \end{pmatrix}$$

La matrice de différence $\Delta_{u_1, u_2}^{(300)}$ est égale à $\Delta_{u_1, u_2}^{(300)} = \begin{pmatrix} 2 & -4 & -1 & -1 \\ 3 & -5 & -1 & -2 \\ 4 & -7 & -1 & -1 \\ 4 & -10 & -2 & -1 \\ 7 & -10 & -2 & -1 \end{pmatrix}$

5.1.2.3 Etude du comportement des collaborateurs

Nous présentons, ici, les résultats obtenus à partir des trois scénarios d'expérimentation utilisateurs présentés précédemment. Dans un premier temps, nous présentons l'analyse comparative effectuée au niveau quantitatif et qualitatif, puis, nous énonçons les principales hypothèses validées.

Analyse quantitative du comportement des collaborateurs. L'intuition qui guide notre analyse repose sur deux hypothèses principales :

1. Les utilisateurs se comportent différemment tout au long de la session.
2. La complémentarité de leurs compétences évolue également.

Pour cela, nous proposons une analyse temporelle des comportements des collaborateurs en modélisant la session de recherche des utilisateurs $S_{u_j}^{(t_i)}$ à partir des caractéristiques $f_k \in F$ estimées à chaque instant t_i de la session, selon des intervalles réguliers de temps d'une minute. Dans ce contexte, nous faisons l'hypothèse que des collaborateurs qui organisent leurs stratégies de recherche sont caractérisés par des comportements différents et opposés vis-à-vis de deux caractéristiques f_k et $f_{k'}$. Une analyse naïve permettant de mettre en exergue de tels comportements repose sur l'estimation des corrélations $\rho(S_{u_1}^{(t_i)}(f_k), S_{u_2}^{(t_i)}(f_{k'}))$ des valeurs des paires de caractéristiques f_k et $f_{k'}$, deux à deux. Cependant, nous favorisons une estimation des corrélations $\rho(\Delta_{u_1, u_2}^{(t_i)}(t_{l'}, f_k), \Delta_{u_1, u_2}^{(t_i)}(t_{l'}, f_{k'}))$ sur la base des différences $\Delta_{u_1, u_2}^{(t_i)}(t_{l'}, f_k)$ et $\Delta_{u_1, u_2}^{(t_i)}(t_{l'}, f_{k'})$ entre collaborateurs u_1 et u_2 vis-à-vis des caractéristiques de comportement f_k et $f_{k'}$, plutôt que des valeurs intrinsèques, afin de mettre en évidence (a) les compétences pour lesquelles un collaborateur est le plus efficace et (b) les compétences pour lesquelles il est plus efficace que son collaborateur. En effet, si l'on considère, par exemple, deux caractéristiques de comportement liées au nombre de requêtes soumises (N_q) et au nombre de pages visitées par requêtes (N_{pq}), savoir que les deux collaborateurs sont différents vis-à-vis de ces caractéristiques n'est pas suffisant. En effet, un utilisateur peut avoir des valeurs plus élevées pour les deux caractéristiques et dans ce cas là, le groupe de collaboration ne dispose pas de compétences complémentaires qu'il peut exploiter parmi les participants. Une analyse des différences, et plus particulièrement des corrélations des différences permet donc de détecter quelle est la meilleure compétence de chacun des collaborateurs complémentaires par rapport aux autres membres du groupe qui peut bénéficier à la collaboration.

Un aperçu général des corrélations des différences des caractéristiques de comportement entre une paire de participants est donné dans la Figure 5.4. Afin de répondre à notre objectif d'identification des complémentarités, nous avons représenté seulement les corrélations négatives. Respectivement, chaque figure représente les corrélations négatives significatives entre paires de caractéristiques de comportement compte tenu des scénarios GS, PM et SansRole. Ces corrélations sont estimées toutes les 5 minutes afin d'identifier l'évolution du com-

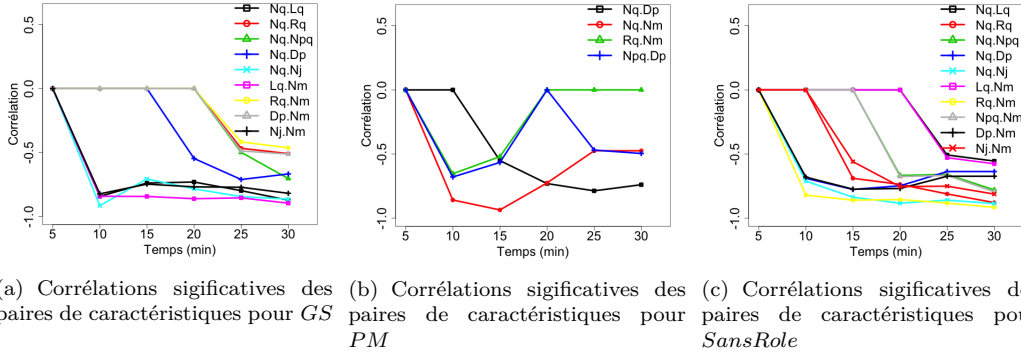


FIGURE 5.4: Analyse des corrélations des caractéristiques de comportement des collaborateurs

portement des collaborateurs durant la session. D'un point de vue global, nous observons que l'ensemble des corrélations, que l'on peut étiqueter comme le rôle implicite des collaborateurs, évolue au fur et à mesure de la session, mettant en évidence l'implication dans la collaboration et la convergence des participants dans la coordination de leurs actions dans la session en terme de compétences de recherche. Nous faisons l'hypothèse que les participants parviennent à identifier leurs similarités et leurs complémentarités lorsque les tendances des courbes de corrélations sont stables par rapport à l'indicateur temporel. A partir des figures, nous pouvons formuler les trois observations suivantes :

1. Les caractéristiques convergent tardivement vers des corrélations significatives (positives ou négatives) pour le GS. En effet, seulement 55% (5/9) des caractéristiques atteignent une corrélation négative à partir de la 20^{ième} minute tandis que la totalité des caractéristiques ne convergent qu'au bout de 25 minutes.
2. 50% (2/4) des caractéristiques impliquées dans le scénario PM varient entre des corrélations positives/négatives/nulles tout au long de la session. Par exemple, la corrélation des différences entre participants des caractéristiques du nombre de pages par requêtes N_{pq} et de durée de visite sur une page D_p n'est pas significative en début de session, négative entre la 10^{ième} et 20^{ième} minute, non significative entre 20 et 25 minutes et enfin négative jusqu'à la fin de la session.
3. Les collaborateurs du scénario sans contrainte de rôles sont caractérisés par des corrélations qui convergent plus rapidement dans la session. Plus particulièrement, nous observons que les collaborateurs atteignent une convergence dans la complémentarité de leur comportement pour 80% (8/10) des caractéristiques 15 minutes après le début de la session.

Dans l'objectif d'effectuer une analyse plus en détail, nous nous concentrons sur les différences entre le scénario sans rôle, noté *SansRole*, et les scénarios PM et

GS où la contrainte des rôles est imposée. Plus particulièrement, notre objectif est de mettre en évidence les couples de caractéristiques de comportement qui se distinguent entre les scénarios. Ainsi, nous nous focalisons sur les paires de caractéristiques qui sont significativement corrélées pour le scénario *SansRole* et qui ne le sont pas pour les scénarios GS et PM respectivement. Bien qu'il n'existe pas de différence significative entre le scénario *GS* et *SansRole*, nous soulignons deux aspects importants qui différencient le scénario *PM* du scénario *SansRole* :

1. Seulement 4 paires de caractéristiques sont recensées dans le scénario *PM* alors que nous en avons identifié 11 pour le scénario *SansRole*. Cela peut être expliqué par le fait que le scénario *SansRole* autorise plus de liberté aux collaborateurs dans la structuration de la collaboration, permettant ainsi de dégager des complémentarités de compétences. A l'inverse, la contrainte des rôles imposée aux utilisateurs n'optimise pas totalement la collaboration, comme indiqué dans la Figure 5.4, dans la mesure où ils peuvent ne pas correspondre au comportement des collaborateurs.
2. Les paires de caractéristiques qui ne sont pas significativement corrélées pour le scénario *PM* mais qui le sont pour le scénario *SansRole* sont principalement des caractéristiques reliées à la paire de rôle Prospecteur-Mineur dans la mesure où le Mineur a pour objectif d'identifier des documents pertinents (N_j) et le Prospecteur doit reformuler des requêtes (N_q) avec un taux de recouvrement faible (R_q) dans l'objectif de favoriser la diversité thématique. Ce sont ces caractéristiques, et particulièrement celles reliées aux requêtes qui ne sont pas pleinement identifiées comme négativement corrélées pour le scénario *PM*. Ce constat renforce notre intuition qu'une inadéquation des rôles par rapport aux compétences des collaborateurs peut mener à des scénarios trop contraignants dans l'organisation de la collaboration. A l'inverse, les collaborateurs impliqués dans des scénarios libres identifient naturellement leurs similarités et leurs complémentarités en termes de compétences de recherche.

Ces observations nous permettent de formuler l'hypothèse que les collaborateurs contraints par les rôles semblent avoir des difficultés à atteindre une convergence en terme de coordination de leurs compétences complémentaires, plus particulièrement pour le scénario PM dans lequel les collaborateurs semblent en plus agir selon des rôles implicites différents tout au long de la session. Ces résultats confirment notre intuition que les participants n'ayant pas de contraintes de rôles imposées semblent mieux se coordonner, et de façon stable, tout au long de la session dans l'objectif d'assurer la complémentarité de leurs compétences vis-à-vis de l'objectif commun de la collaboration et de la variabilité de leur comportement en réponse aux interactions durant la session.

Analyse qualitative du comportement des collaborateurs. Afin de compléter cette étude quantitative, nous souhaitons analyser ces aspects d'un point de

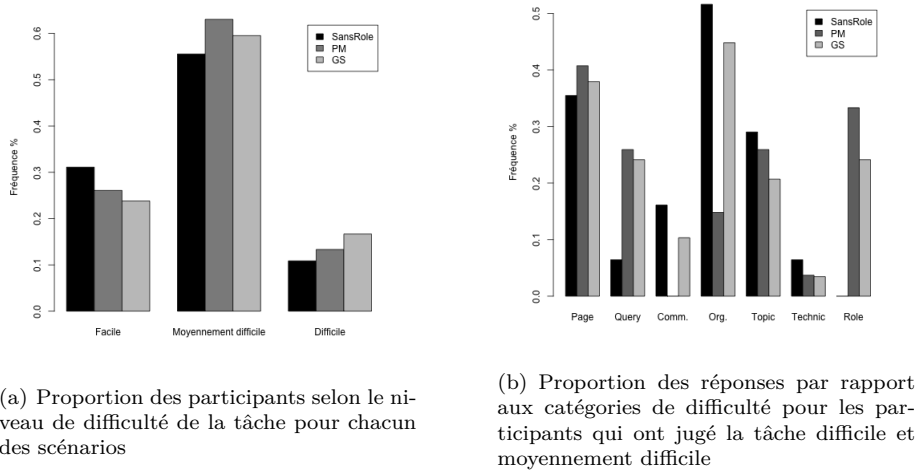


FIGURE 5.5: Analyse quantitative de la difficulté des tâches

vue qualitative grâce à l’exploitation des questionnaires post-tâche. À l’aide de la question liée au niveau de difficulté de la tâche issue du questionnaire post-tâche, nous avons représenté dans la Figure 5.5(a) le nombre de participants, impliqués dans chacun des scénarios, qui ont jugé la tâche facile, moyennement difficile et difficile. Comme nous pouvons l’observer, il n’y a pas de différence importante en terme de difficulté de tâche entre les scénarios. Afin de compléter cette analyse, nous avons étudié les réponses de la question liée au type de difficulté du questionnaire post-tâche dont la codification manuelle des réponses est reportée dans le Tableau 5.1. Ainsi, nous représentons dans la Figure 5.5(b) les principales catégories de difficultés particulièrement pour les utilisateurs qui ont trouvé la tâche moyennement difficile et difficile. Nous avons trouvé que les principales raisons qui expliquent la difficulté de la tâche sont liées à la communication (*Comm*), à l’organisation (*Org*) et aux rôles (*Role*). Nous pouvons regrouper ces difficultés comme des difficultés liées à la coordination de la tâche. Comme illustré dans la Figure 5.5(b), nous établissons que (a) la coordination liée à la communication et à l’organisation est relativement facile dans le scénario PM et (b) l’identification des rôles implicites dans un scénario *SansRole* est relativement facile tandis que pour un scénario avec rôles, PM et GS, elle semble plus difficile. En effet, nous avons indentifié quatre participants impliqués dans ces scénarios qui ont affirmé avoir des difficultés à suivre les rôles : “*Ce que j’ai trouvé difficile était de suivre le rôle*”, “*Me contraindre au rôle*”. Ces observations confirment nos précédentes conclusions issues des analyses quantitatives basées sur les corrélations. Nous avons également observé d’autres types de difficultés liées aux catégories *Page*, *Topic* et *Technic*.

Discussion. Dans cette analyse, nous avons comparé le comportement des collaborateurs suivant deux scénarios principaux distingués selon (a) une attribution des paires de rôles prédéfinis (Prospecteur-Mineur ou Rassembleur-Surveilleur) ou (b) une recherche libre où les utilisateurs laissent émerger de façon implicite leur propre rôle. D’après les résultats, nous pouvons déduire et/ou confirmer les hypothèses suivantes :

1. Les utilisateurs se comportent différemment sur l’ensemble de la session dans la mesure où ils n’utilisent pas les mêmes compétences tout au long de la session. Une convergence des compétences semble être atteinte en fin de session, mais un intervalle de temps semble nécessaire aux collaborateurs impliqués dans les sessions PM et GS pour se coordonner et identifier les compétences pour lesquelles chacun est le plus efficace.
2. Les scénarios basés sur des rôles fixes semblent soulever deux limites : (a) la convergence des complémentarités est tardive dans la session et (b) les collaborateurs semblent avoir des difficultés à appliquer les spécificités de la paire de rôles.
3. Les collaborateurs qui ne sont pas guidés par des rôles prédéfinis semblent avoir davantage de difficulté à structurer la communication durant la session du fait de la latitude plus large de leurs actions.

5.1.3 Contribution et questions de recherche

Compte tenu des conclusions obtenues lors de l’analyse exploratoire, nous proposons de bénéficier des avantages des deux types de médiation en considérant une approche hybride basée à la fois sur une médiation utilisateur et une médiation système qui permet de guider la collaboration en assignant dynamiquement des rôles prédéfinis aux collaborateurs tout en laissant les utilisateurs faire émerger leur propre rôle. Pour cela, nous proposons le scénario de recherche suivant :

1. En début de session, les collaborateurs sont similaires et sont caractérisés par les mêmes comportements.
2. Au fur et à mesure de la session, les utilisateurs sont capables de s’organiser dans la collaboration afin de faire émerger leur rôles et de bénéficier de leurs compétences complémentaires, qui peuvent évoluer au fur et à mesure de la session en fonction des interactions avec les sources d’information ou les membres du groupe de collaboration. Afin de caractériser la complémentarité des compétences, nous considérons le comportement relatif entre les utilisateurs afin d’identifier les compétences pour lesquelles ils sont les plus efficaces ; à l’inverse des travaux basés sur les modèles de comportement des utilisateurs (Agichtein et al., 2006; Leung et al., 2012; Radinsky et al., 2013) qui reposent principalement sur les caractéristiques intrinsèques des utilisateurs.

A partir de ces hypothèses, nous proposons d’apprendre les rôles des collaborateurs en fonction de leurs actions selon un processus dynamique d’apprentissage

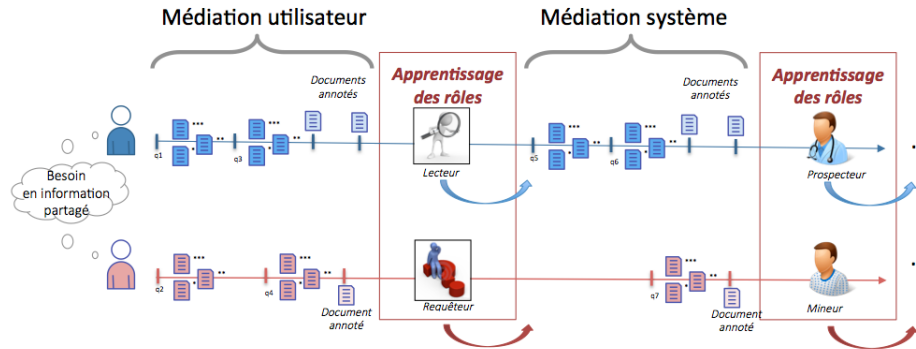


FIGURE 5.6: Contextualisation de notre approche par rapport à la notion de session hybride vis-à-vis de la médiation

des rôles. La Figure 5.6 illustre notre approche et plus particulièrement son intégration dans le contexte d'une session de recherche de médiation hybride. Plus particulièrement, ce modèle étudie les différences de comportement entre collaborateurs afin de leur suggérer des rôles de la littérature (Golovchinsky et al., 2009b; Pickens et al., 2008; Shah et al., 2010) qui permettent d'optimiser la collaboration en adoptant en temps réel les modèles d'ordonnancement des documents associés. Notre contribution repose sur les trois questions de recherche suivantes :

1. *Dans quelle mesure les collaborateurs se comportent-ils différemment au cours de la session ?* La différence entre les utilisateurs est une condition nécessaire à la détection de la complémentarité de leurs compétences. Ainsi, nous proposons ici de modéliser les collaborateurs par des variables de comportement de recherche et d'étudier la significativité de leurs différences vis-à-vis de ces variables.
2. *Comment inférer les rôles des utilisateurs en prenant en compte leurs actions dans la session de RIC ?* Une fois que les collaborateurs ont été identifiés comme différents selon certains aspects de leur comportement, il est donc possible de détecter les complémentarités caractéristiques des collaborateurs afin d'apprendre leur rôle. Nous supposons des rôles prédéfinis qui sont caractérisés par des corrélations négatives entre les variables de comportement afin de mettre en évidence leurs compétences complémentaires.
3. *Comment utiliser les rôles afin d'améliorer la session de RIC et le processus de RI en général ?* A ce stade, l'enjeu est de guider la collaboration afin de bénéficier des compétences des collaborateurs. Ainsi, nous proposons d'ordonner les documents selon le modèle de RIC (Pickens et al., 2008; Shah et al., 2010) associé à la paire de rôles détectée afin de fournir des ordonnancements les plus en adéquation avec les compétences des collaborateurs.

5.2 Notations : modélisation des patrons de rôles

Notre contribution considère une session de RIC synchrone S entre une paire d'utilisateurs (u_1, u_2) qui collaborent pendant une période de temps T afin de satisfaire un besoin en information partagé I . Chaque utilisateur agit individuellement pendant la session S en formulant ses propres requêtes. L'objectif de notre modèle étant de détecter, lors d'une session S , les rôles des utilisateurs (u_1, u_2) parmi un ensemble R de paires de rôles prédéfinis. Chaque paire de rôles est associée à des stratégies de recherche particulières que nous avons choisi de modéliser sous la forme d'un patron de rôles. Nous définissons les rôles des collaborateurs, dans ce qui suit.

Une paire $R_{1,2}$ de rôles (r_1, r_2) peut être assignée à une paire d'utilisateurs (u_1, u_2) sous l'hypothèse qu'ils aient des compétences complémentaires, modélisées par les caractéristiques de comportement. Ainsi, nous modélisons une paire de rôles par un patron $P^{R_{1,2}} \in P$, où P est l'ensemble des patrons de rôles associé à l'ensemble R des paires de rôles. Le patron $P^{R_{1,2}}$, présenté dans l'illustration 5.2, consiste en un ensemble de deux éléments :

1. Une matrice de caractéristiques $F^{R_{1,2}} \in \mathcal{R}^{F \times F}$ qui modélise le comportement potentiel des collaborateurs s'ils étaient assignés aux rôles (r_1, r_2) . Chaque élément $F^{R_{1,2}}(f_k, f_{k'})$ correspond à la relation attendue entre les caractéristiques de comportement f_k et $f_{k'}$ sous-jacentes à la paire de rôles (r_1, r_2) . Compte tenu de notre objectif de détecter les complémentarités entre collaborateurs, les relations possibles entre deux caractéristiques f_k et $f_{k'}$ correspondent à des corrélations positives (+1), négatives (-1) et nulles (0), traduisant respectivement une similarité, complémentarité ou absence de liens dans les compétences de recherche. Les caractéristiques impliquées dans une corrélation négative constituent le noyau $\mathcal{K}^{R_{1,2}}$ du patron $P^{R_{1,2}}$ de la paire de rôles (r_1, r_2) , modélisé ainsi : $\mathcal{K}^{R_{1,2}} = \{(f_k, f_{k'}); F^{R_{1,2}}(f_k, f_{k'}) = -1\}$.
2. Une fonction d'attribution des rôles $Role(S_{u_1}^{(t_i)}, S_{u_2}^{(t_i)}, R_{1,2})$, qui permet d'assigner à chaque collaborateur u_1 et u_2 un des rôles de la paire $R_{1,2}$ compte tenu de leur comportement de recherche modélisé par la représentation de leur session de recherche individuelle $S_{u_1}^{(t_i)}$ et $S_{u_2}^{(t_i)}$. Dans cet objectif, nous comparons les valeurs $S_{u_1}^{(t_i)}(t_{i'}, f_k)$ et $S_{u_2}^{(t_i)}(t_{i'}, f_k)$ pour au moins une des caractéristiques f_k recensées dans le noyau $\mathcal{K}^{R_{1,2}}$. Cette comparaison permet d'identifier quel rôle de la paire $R_{1,2}$ est le plus vraisemblable pour chaque collaborateur. Le résultat de cette fonction d'attribution consiste en un tuple de rôles ordonnés (r, r') dont le premier élément correspond au rôle assigné au collaborateur u_1 et, par analogie, le deuxième est le rôle attribué au collaborateur u_2 . Ainsi, nous définissons la fonction d'attribution $Role(S_{u_1}^{(t_i)}, S_{u_2}^{(t_i)}, R_{1,2})$ de la paire

de rôles $R_{1,2}$ aux collaborateurs u_1 et u_2 de la façon suivante :

$$\begin{aligned} Role : (S_{u_1}^{(t_l)}, S_{u_2}^{(t_l)}, R_{1,2}) &\rightarrow \{(r_1, r_2), (r_2, r_1)\} \\ (S_{u_1}^{(t_l)}, S_{u_2}^{(t_l)}, R_{1,2}) &\mapsto \begin{cases} (r_1, r_2) & \text{si } S_{u_1}^{(t_l)}(t_{l'}, f_k) > S_{u_2}^{(t_l)}(t_{l'}, f_k) \\ (r_2, r_1) & \text{sinon} \end{cases} \end{aligned} \quad (5.2)$$

Illustration 5.2 Patron de rôles.

Imaginons une paire de rôles de “lecteur” et de “requêteur” dont l’objectif est respectivement (a) de lire un grand nombre de documents afin d’obtenir une connaissance profonde du besoin en information et (b) de soumettre un grand nombre de requêtes pour une analyse surfacique des résultats de recherche, sans spécialement visiter les pages et en se contentant de visualiser le titre et les méta-données des pages. Nous pouvons donc construire le patron de rôles sur la base des caractéristiques du nombre de requêtes N_q et de pages N_p en formulant une corrélation négative entre ces deux caractéristiques. On obtient donc $P^{R_{lec, req}}$ composé de deux éléments :

- une matrice de caractéristiques $F^{R_{lec, req}} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$

avec un noyau $\mathcal{K}^{R_{lec, req}} = \{(N_q, N_p)\}$

- une fonction d’attribution

$$\begin{aligned} Role : (S_{u_1}^{(t_l)}, S_{u_2}^{(t_l)}, R_{lec, req}) &\rightarrow \{(lecteur, requêteur), (requêteur, lecteur)\} \\ (S_{u_1}^{(t_l)}, S_{u_2}^{(t_l)}, R_{lec, req}) &\mapsto \begin{cases} (lecteur, requêteur) & \text{si } S_{u_1}^{(t_l)}(t_{l'}, N_p) > S_{u_2}^{(t_l)}(t_{l'}, N_p) \\ (requêteur, lecteur) & \text{sinon} \end{cases} \end{aligned}$$

5.3 Modèle hybride de RIC basé sur l’apprentissage des rôles des collaborateurs

Dans cette section, nous présentons notre modèle d’apprentissage de rôles pré-définis pour un contexte de RIC impliquant une paire de collaborateurs. Afin de répondre à notre hypothèse de variabilité des comportements des collaborateurs, ce modèle est dynamique et peut être appliqué à tout instant t_l de la session. Comme illustré dans la Figure 5.7 et détaillé dans l’Algorithme 3, notre modèle repose sur deux étapes :

1. L’identification des différences de comportement entre collaborateurs qui permet de déterminer si les collaborateurs sont potentiellement complémentaires en terme de comportement.
2. L’identification des rôles de collaboration à partir de l’inférence des complémentarités des comportements afin de les réinjecter dans un modèle d’ordonnement collaboratif.

L'objectif, à l'issue de cette méthodologie, est d'adapter l'ordonnancement des documents en fonction des rôles détectés en appliquant, en temps réel, le modèle de RIC associé.

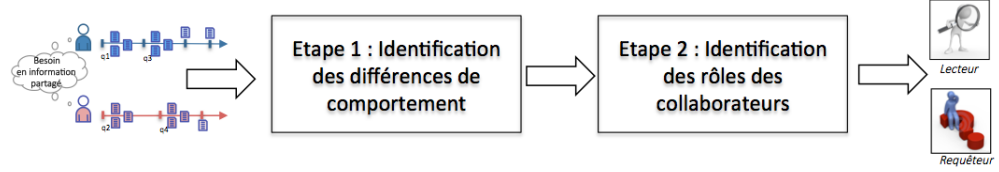


FIGURE 5.7: Etape du modèle d'apprentissage des rôles

Algorithm 3: Apprentissage des rôles en contexte de RIC

Données: $S_{u_1}^{(t_1)}, S_{u_2}^{(t_1)}, P$

Résultat: $\{(r_1, r_2), (r_2, r_1)\} \in R$

begin

/* Etape 1: Identification des différences de comportement entre collaborateurs */

Pour chaque caractéristique $f_k \in F$ **faire**

$p_{u_1, u_2}^{(t_1)}(f_k) = \text{ks.test}(S_{u_1}^{(t_1)}(f_k), S_{u_2}^{(t_1)}(f_k))$

/* Etape 2: Identification des rôles des collaborateurs */

Pour chaque caractéristique $f_k \in F$ **faire**

Pour chaque caractéristique $f_{k'} \in F$ **faire**

Si $(p_{u_1, u_2}^{(t_1)}(f_k) < \theta)$ **et** $(p_{u_1, u_2}^{(t_1)}(f_{k'}) < \theta)$ **alors**

$C_{u_1, u_2}^{(t_1)}(f_k, f_{k'}) = \rho(\Delta_{u_1, u_2}^{(t_1)}(f_k), \Delta_{u_1, u_2}^{(t_1)}(f_{k'}))$

sinon

$C_{u_1, u_2}^{(t_1)}(f_k, f_{k'}) = 0$

$R_{1,2}^* = \underset{\text{sous la contrainte :}}{\text{argmin}}_{R_{1,2}} \|F^{R_{1,2}} \ominus C_{u_1, u_2}^{(t_1)}\|$

$\forall (f_k, f_{k'}) \in \mathcal{K}^{R_{1,2}} \quad F^{R_{1,2}}(f_k, f_{k'}) - C_{u_1, u_2}^{(t_1)}(f_k, f_{k'}) > -1$

Si $R_{1,2}^* \neq \text{null}$ **alors**

Retourne $\text{Role}(S_{u_1}^{(t_1)}, S_{u_2}^{(t_1)}, R_{1,2}^*)$

Retourne null

5.3.1 Identification des différences de comportement entre collaborateurs

Cette étape découle des conditions de complémentarité des compétences entre collaborateurs et a pour objectif de déterminer, à partir de leurs actions, pour

quelles caractéristiques de comportement ils sont différents. Plus formellement, l'objectif est de déterminer le niveau de différence entre les deux collaborateurs u_1 et u_2 vis-à-vis de chaque caractéristique de comportement $f_k \in F$. Pour cela, nous utilisons le test statistique non paramétrique de Kolmogorov-Smirnov, noté *ks.test()* dans l'Algorithme 3, entre les deux vecteurs $S_{u_1}^{(t_l)}(f_k)$ et $S_{u_2}^{(t_l)}(f_k)$ qui modélisent la session de recherche des collaborateurs u_1 et u_2 respectivement. La valeur de significativité du test, appelée *p-value*, est notée $p_{u_1, u_2}^{(t_l)}(f_k)$ et, lorsqu'elle est inférieure à θ (généralement $\theta = 0.05$), correspond à une différence significative qui implique que les collaborateurs u_1 et u_2 ont des différences dans leurs compétences de recherche ou leurs stratégies qui, par conséquent, peuvent impliquer des rôles différents dans la session. Un exemple est présenté dans l'Illustration 5.3.

Illustration 5.3 Différence de comportement entre collaborateurs.

Nous souhaitons apprendre les rôles des collaborateurs à la cinquième minute de la session, soit $t_l = 300$. Notre première étape consiste en l'identification des différences de comportement entre collaborateurs. En appliquant le test de Kolmogorov-Smirnov sur les matrices représentant les sessions utilisateurs, présentées dans l'Illustration 5.1, nous obtenons les valeurs de significativité suivantes :

- nombre de requêtes : $p_{u_1, u_2}^{(t_l)}(N_q) = 0.01348$
- nombre de pages : $p_{u_1, u_2}^{(t_l)}(N_d) = 0.01348$
- nombre de commentaires : $p_{u_1, u_2}^{(t_l)}(N_c) = 0.01348$
- nombre de notes : $p_{u_1, u_2}^{(t_l)}(N_n) = 0.08152$

Nous pouvons ainsi dire que les collaborateurs sont différents vis-à-vis de trois caractéristiques de comportement qui font référence au nombre de requêtes, de pages et de commentaires. Ils ne sont cependant pas significativement différents par rapport au nombre de notes.

5.3.2 Identification des rôles des collaborateurs

Contrairement à la première étape qui souligne les différences entre collaborateurs, l'objectif de cette étape est de caractériser ces différences de comportement afin de mettre en évidence les stratégies/compétences de recherche complémentaires des collaborateurs et de détecter la paire de rôles qui leur correspond le plus. Dans cet objectif, notre intuition est d'estimer la corrélation entre paires de caractéristiques de comportement. Cependant, dans l'objectif d'éviter une assignation a priori des rôles pour chacun des collaborateurs et d'identifier les compétences de recherche pour lesquelles ils sont les plus efficaces, nous estimons les corrélations sur les vecteurs de différences des caractéristiques, paires à paires. La corrélation est ainsi calculée entre les deux vecteurs de différences entre collaborateurs vis-à-vis des caractéristiques de comportement $\Delta_{u_1, u_2}^{(t_l)}(f_k)$ et $\Delta_{u_1, u_2}^{(t_l)}(f_{k'})$ avec $(f_k, f_{k'}) \in F \times F$ sous la condition que chacune des deux

caractéristiques impliquées ait été détectées comme significativement différente par le test de Kolmogorov-Smirnov entre les deux collaborateurs. Dans le cas contraire, la corrélation reçoit la valeur de 0. Au final, la matrice de corrélations $C_{u_1, u_2}^{(t_l)} \in \mathcal{R}^{F \times F}$ est construite de la façon suivante (Illustration 5.4) :

$$C_{u_1, u_2}^{(t_l)}(f_k, f_{k'}) = \begin{cases} \rho(\Delta_{u_1, u_2}^{(t_l)}(f_k), \Delta_{u_1, u_2}^{(t_l)}(f_{k'})) & \text{si } ((p_{u_1, u_2}^{(t_l)}(f_k) < \theta) \\ & \text{et } (p_{u_1, u_2}^{(t_l)}(f_{k'}) < \theta)) \\ 0 & \text{sinon} \end{cases} \quad (5.3)$$

où $\rho(\Delta_{u_1, u_2}^{(t_l)}(f_k), \Delta_{u_1, u_2}^{(t_l)}(f_{k'}))$ représente la corrélation entre les deux vecteurs de différences $\Delta_{u_1, u_2}^{(t_l)}(f_k)$ et $\Delta_{u_1, u_2}^{(t_l)}(f_{k'})$ vis-à-vis des caractéristiques f_k et $f_{k'}$. θ correspond au niveau de significativité du test de Kolmogorov-Smirnov et $p_{u_1, u_2}^{(t_l)}(f_k)$ représente la p-value associée au test pour la caractéristique f_k .

Illustration 5.4 Extraction de la matrice de corrélation.

A partir des illustrations précédentes, nous pouvons construire la matrice $\Delta_{u_1, u_2}^{(300)}$ des différences entre collaborateurs. Nous obtenons :

$$\Delta_{u_1, u_2}^{(300)} = \begin{pmatrix} 2 & -4 & -1 & -1 \\ 3 & -5 & -1 & -2 \\ 4 & -7 & -1 & -1 \\ 4 & -10 & -2 & -1 \\ 7 & -10 & -2 & -1 \end{pmatrix}$$

où les colonnes représentent les caractéristiques de comportement $\{N_q, N_d, N_c, N_n\}$ et les lignes les intervalles de temps séparés de 60 secondes.

A partir de cette matrice, nous pouvons extraire la matrice $C_{u_1, u_2}^{(300)}$ des corrélations des différences entre paires de caractéristiques de comportement. Nous obtenons ainsi la matrice suivante :

$$C_{u_1, u_2}^{(300)} = \begin{pmatrix} 1 & -0.8186713 & -0.731925 & 0 \\ -0.8186713 & 1 & 0.9211324 & 0 \\ -0.731925 & 0.9211324 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Nous observons que les paires de caractéristiques N_q/N_p et N_q/N_c sont négativement corrélées alors que les caractéristiques N_p/N_n sont positivement corrélées. La dernière ligne et la dernière colonne ont des valeurs nulles car les collaborateurs ne sont pas significativement différents pour la caractéristique N_n .

Ensuite, le patron de rôles est détecté si la matrice de comportement des collaborateurs respecte deux conditions :

1. Les paires de caractéristiques impliquées dans le noyau du patron de rôles doivent être significativement différentes et corrélées négativement pour la paire de collaborateurs. En effet, nous faisons l'hypothèse que, pour être assignés à une paire de rôles, les collaborateurs doivent au moins se comporter significativement différemment vis-à-vis des caractéristiques impliquées dans le noyau du patron de rôles sous-jacent. Cette condition peut être formulée par le fait que si la mesure de corrélation $F^{R_{1,2}}(f_k, f_{k'})$ reliée au patron de rôles P^{r_1, r_2} est égale à -1 , alors la corrélation $C_{u_1, u_2}^{(t_l)}(f_k, f_{k'})$ vis-à-vis des différences des caractéristiques f_k et $f_{k'}$ entre collaborateurs doit être strictement inférieure à 0.
2. La matrice de corrélations du patron de rôles doit être la plus similaire à la matrice de corrélations reflétant le comportement des collaborateurs.
Nous souhaitons savoir si une de ces deux paires de rôles peut être attribuée au couple (u_1, u_2) . Nous sommes conscients qu'un appariement exact entre ces deux matrices de corrélations est difficile à atteindre compte tenu qu'une complémentarité des compétences dans le patron de rôles est notée par une corrélation négative égale à 1. Ainsi, nous proposons d'identifier la paire de rôles selon un appariement approximé où la matrice de corrélations $F^{R_{1,2}}$ du patron de rôles doit être la plus similaire à celle $C_{u_1, u_2}^{(t_l)}$ obtenue pour la paire (u_1, u_2) de collaborateurs à l'instant t_l .

Plus formellement, une paire $R_{1,2}$ de rôles (r_1, r_2) est assignée à une paire d'utilisateurs (u_1, u_2) à l'instant t_l si elle satisfait les conditions suivantes :

$$\operatorname{argmin}_{R_{1,2}} \|F^{R_{1,2}} \ominus C_{u_1, u_2}^{(t_l)}\| \quad (5.4)$$

sous la contrainte :

$$\forall_{(f_k, f_{k'}) \in \mathcal{K}^{R_{1,2}}} F^{R_{1,2}}(f_k, f_{k'}) - C_{u_1, u_2}^{(t_l)}(f_k, f_{k'}) > -1$$

où $\|\cdot\|$ représente la norme de Frobenius et \ominus est l'opérateur de soustraction entre deux éléments de chaque matrice défini ainsi :

$$F^{R_{1,2}}(f_k, f_{k'}) \ominus C_{u_1, u_2}^{(t_l)}(f_k, f_{k'}) = \begin{cases} F^{R_{1,2}}(f_k, f_{k'}) - C_{u_1, u_2}^{(t_l)}(f_k, f_{k'}) \\ \quad \text{si } F^{R_{1,2}}(f_k, f_{k'}) \in \{-1; 1\} \\ 0 \quad \text{sinon} \end{cases} \quad (5.5)$$

Une fois que la paire $R_{1,2}$ de rôles a été détectée pour une paire de collaborateurs u_1 et u_2 , les rôles sont assignés aux collaborateurs selon la fonction d'attribution des rôles $\operatorname{Role}(S_{u_1}^{(t_l)}, S_{u_2}^{(t_l)}, R_{1,2})$, définie dans l'Equation 5.2.

L'Illustration 5.5 présente l'exemple associé à cette étape.

Illustration 5.5 Identification des rôles des collaborateurs.

Considérons les deux patrons de rôles suivants :

1. Les rôles de lecteur et requêteur définis précédemment où $P^{R_{lec,req}}$ avec :

$$F^{R_{lec,req}} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \mathcal{K}^{R_{lec,req}} = \{(N_q, N_p)\}$$

$$Role : (S_{u_1}^{(t_l)}, S_{u_2}^{(t_l)}, R_{lec,req}) \rightarrow \{(lecteur, requêteur), (requêteur, lecteur)\}$$

$$(S_{u_1}^{(t_l)}, S_{u_2}^{(t_l)}, R_{lec,req}) \mapsto \begin{cases} (lecteur, requêteur) & \text{si } S_{u_1}^{(t_l)}(t_l, N_p) > S_{u_2}^{(t_l)}(t_l, N_p) \\ (requêteur, lecteur) & \text{sinon} \end{cases}$$

2. Les rôles de juge et requêteur caractérisés respectivement par leur capacité à juger les documents pertinents et à soumettre des requêtes où $P^{R_{juge,req}}$ avec :

$$F^{R_{juge,req}} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \mathcal{K}^{R_{juge,req}} = \{(N_q, N_c)\}$$

$$Role : (S_{u_1}^{(t_l)}, S_{u_2}^{(t_l)}, R_{juge,req}) \rightarrow \{(juge, requêteur), (requêteur, juge)\}$$

$$(S_{u_1}^{(t_l)}, S_{u_2}^{(t_l)}, R_{juge,req}) \mapsto \begin{cases} (juge, requêteur) & \text{si } S_{u_1}^{(t_l)}(t_l, N_c) > S_{u_2}^{(t_l)}(t_l, N_c) \\ (requêteur, juge) & \text{sinon} \end{cases}$$

Pour cela, nous vérifions avant tout que les caractéristiques impliquées dans les noyaux \mathcal{K} des paires de rôles ont été détectées comme significativement différentes pour la paire de collaborateurs. Ce qui est le cas pour les deux paires de rôles considérées. Nous pouvons donc appliquer la formule présentée dans l'Equation 5.4 :

1. Paire de lecteur et de requêteur :

$$\begin{aligned} \|F^{R_{lec,req}} \ominus C_{u_1, u_2}^{(300)}\| &= \begin{pmatrix} 0 & -1 - (-0.8186713) \\ -1 - (-0.8186713) & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0.183287 \\ 0.183287 & 0 \end{pmatrix} \end{aligned}$$

La norme de Frobenius de cette matrice est égale à $\sqrt{0.183287^2} = 0.183287$.

2. Paire de juge et de requêteur :

$$\begin{aligned} \|F^{R_{juge,req}} \ominus C_{u_1, u_2}^{(300)}\| &= \begin{pmatrix} 0 & -1 - (-0.731925) \\ -1 - (-0.731925) & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0.268174 \\ 0.268174 & 0 \end{pmatrix} \end{aligned}$$

La norme de Frobenius de cette matrice est égale à $\sqrt{0.268174^2} = 0.268174$.

La paire de rôles qui minimise la norme est la paire lecteur/requêteur. Cette paire de rôles est donc assignée aux collaborateurs u_1 et u_2 . Nous souhaitons désormais identifier plus particulièrement quel rôle de la paire est associé à chaque utilisateur. Pour cela, nous utilisons la fonction d'attribution $Role : (S_{u_1}^{(t_l)}, S_{u_2}^{(t_l)}, R_{lec,req})$ et observons que le rôle de lecteur est assigné à l'utilisateur qui a la valeur la plus élevée de la caractéristique N_p . Par conséquent, u_1 est labélisé comme requêteur et u_2 est associé au rôle de lecteur.

5.4 Expérimentation et résultats

Notre protocole d'évaluation est de type *log study* et repose sur les logs de deux expérimentations utilisateurs. Les objectifs qui ont guidé notre expérimentation sont les suivants :

1. Identifier les différences entre utilisateurs afin de démontrer la validité de l'intuition de notre contribution qui consiste en l'apprentissage des rôles sur la base des différences entre utilisateurs.
2. Evaluer l'impact de notre modèle d'apprentissage des rôles sur l'efficacité d'une session de recherche d'information collaborative.

Dans ce qui suit, nous présentons dans un premier temps les jeux de données ainsi que les traitements réalisés pour les construire. Ensuite, nous énonçons et discutons les résultats obtenus.

5.4.1 Construction des jeux de données

5.4.1.1 Expérimentations utilisateurs

Dans notre évaluation, nous avons utilisé deux expérimentations utilisateurs, différentes de celles présentées dans l'analyse préliminaire afin d'éviter un biais dans les résultats. Elles ont cependant été réalisées dans les mêmes conditions avec des participants, recrutés quant à eux dans l'université de Rutgers aux Etats-Unis, dont la tâche était d'identifier le maximum d'informations en rapport avec un sujet exploratoire afin d'écrire un rapport. Les participants disposaient du plugin Firefox proposé par Coagmento (Shah and González-Ibáñez, 2011a), qui regroupe les mêmes fonctionnalités (gestion de l'information, communication et sauvegarde des actions des collaborateurs) que dans le prototype que nous avons développé. Des statistiques sur les deux expérimentations utilisateurs sont décrites dans le tableau 5.4.

	US1	US2
Sujet	Gulf oil spill	Global warming
Nombre de paires	60	10
Nombre total de pages web visitées	4326	1935
Nombre total de pages web sauvegardées/notées	254	-
Nombre total de snippets sauvegardés	485	208
Nombre total de requêtes soumises	921	313
Nombre moyen de mots par requête	5.45	4.73

TABLE 5.4: Données statistiques des expérimentations utilisateurs

Dans ce qui suit, nous spécifions les caractéristiques intrinsèques à chacune des expérimentations utilisateurs considérées dans notre évaluation.

US1. Cette expérimentation implique 60 paires de collaborateurs, non guidés par des rôles, qui recherchent des informations sur le sujet “*Gulf oil spill*”. La session de recherche collaborative dure 25 minutes et les participants ont été assignés, de façon aléatoire à une collaboration de proximité ou éloignée.

US2. Cette expérimentation implique 10 groupes de participants, non guidés par des rôles, distincts de la première expérimentation. La configuration se différencie de l’expérimentation US1 par (a) la tâche exploratoire qui traite du sujet “*global warming*”; (b) la durée de la tâche de 30 minutes et (4) la collaboration qui était de proximité.

5.4.1.2 Collection de documents

Dans l’objectif d’ordonner les documents en fonction des rôles grâce à notre médiation système, nous avons construit un seul index de documents pour les deux expérimentations à partir des sources d’évidence suivantes :

1. Les pages visitées par l’ensemble des participants des deux expérimentations.
2. Les 100 premières pages résultats issues du moteur de recherche public Google en réponse aux requêtes soumises par les participants tout au long de l’expérimentation. Notons que les résultats de Google ont été extraits par l’API Google Custom Search³.

L’ensemble des pages extraites a été parsé afin d’extraire le contenu textuel inscrit entre les balises <title> et <p>. Au final, nous avons construit une collection de 74 844 pages web.

5.4.1.3 Caractéristiques de comportement

A partir des données des expérimentations, nous avons enrichi les trois catégories de caractéristiques de comportement présentées dans le Tableau 5.3 lors de l’analyse préliminaire afin de modéliser les actions de recherche des participants (cf. Tableau 5.5) :

1. La première catégorie, notée C_R , analyse le processus de soumission des requêtes de chacun des participants. Nous nous sommes intéressés à identifier si un participant avait l’habitude de formuler un grand nombre de requêtes (N_q), ou également s’il avait des facilités à favoriser la diversité des résultats en estimant le recouvrement entre requêtes successives (R_q) basé sur le vocabulaire utilisé dans les requêtes. Enfin, nous avons considéré des caractéristiques basées sur la longueur des requêtes (L_q) et leur succès (S_q) mesuré par le nombre de pages visitées plus de 30 secondes en moyenne par requête. Cette hypothèse de succès a été formulée par Fox et al. (2005) et White et al. (2010) qui ont inféré du comportement des utilisateurs qu’une page est potentiellement intéressante si le temps passé sur cette page est supérieur à 30 secondes.

3. <https://developers.google.com/custom-search/>

		Description	Mesure
C_R	N_q	Nombre de requêtes	Nombre de requêtes soumises
	L_q	Longueur des requêtes	Nombre moyen de mots par requête
	S_q	Succès des requêtes	Ratio moyen de pages visitées plus de 30 secondes par rapport au nombre de requêtes soumises
	R_q	Recouvrement entre requêtes	Ratio moyen de mots communs entre deux requêtes successives
C_P	N_p	Nombre de pages	Nombre de pages visitées
	N_{pq}	Nombre de pages par requête	Nombre moyen de pages visitées par requêtes
	D_p	Durée de visite	Durée moyenne en seconde d'une visite sur une page
C_S	N_s	Nombre de snippets	Nombre de snippets
	N_{sq}	Nombre de snippets par requêtes	Nombre moyen de snippets par requêtes

TABLE 5.5: Caractéristiques de comportement pour modéliser la session

2. La seconde catégorie, notée C_P , analyse le comportement des participants relatif aux pages web. Pour cela, nous considérons, par exemple, les caractéristiques de comportement liées au temps passé sur les pages web (D_p), ainsi qu'au nombre de pages visitées (N_p), éventuellement normalisé par le nombre de requêtes soumises (N_{pq}), qui montre si un participant est habitué à lire beaucoup de documents dans une session de recherche.
3. La troisième catégorie, notée C_S , analyse les actions de sauvegarde des participants au travers des snippets, particulièrement, car cette fonctionnalité a été utilisée dans les deux expérimentations. Ainsi, le nombre de snippets, qu'il soit normalisé par le nombre de requêtes soumises (N_{sq}) ou non (N_s), mesure la capacité d'un participant à identifier des documents intéressants.

Lors de l'exécution de notre modèle d'apprentissage des rôles pour un instant t_l , notons que nous avons estimé les valeurs de ces caractéristiques en prenant en compte une fenêtre temporelle $[0..t_l]$.

5.4.1.4 Les rôles

Compte tenu des approches de RIC proposées dans l'état de l'art, nous considérons les rôles de Rassembleur-Surveilleur (Shah et al., 2010) et de Prospecteur-Mineur (Pickens et al., 2008) pour lesquels un modèle d'ordonnancement a été proposé. Pour chacun de ces rôles, nous expliquons l'intuition qui nous a permis de construire les patrons de rôles associés, nécessaires à notre méthodologie d'apprentissage des rôles.

1. Les rôles de Rassembleur –*Gatherer*– et de Surveilleur –*Surveyor*– (GS). A partir de leur définition, nous pouvons énoncer que leur principale différence est basée sur le fait que le Rassembleur a pour objectif d'identifier les documents les plus pertinents tandis que le Surveilleur passe moins de temps sur les pages afin de capturer rapidement les différents aspects d'un besoin en information. Nous pouvons donc déduire que (a) le Rassembleur aurait tendance à avoir un taux de recouvrement entre les requêtes plus élevé que le Surveilleur dans la mesure où ses recherches se focalisent sur le coeur du sujet et (b) la durée de visite et le succès d'une requête est plus faible pour le Rassembleur du fait de son analyse surfacique des documents. Par conséquent, nous définissons le patron $P^{R_{GS}}$ des rôles de Rassembleur, noté Ga , et de Surveilleur, noté Su , selon des corrélations négatives entre ces trois variables $\{R_q, D_p, S_q\}$ de la façon suivante :

$$\begin{aligned}
P^{R_{GS}} &= \{F^{R_{GS}}, Role(S_{u_1}^{(t_l)}, S_{u_2}^{(t_l)}, R_{GS})\} \quad (5.6) \\
avec F^{R_{GS}} &= \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \\
et Role(S_{u_1}^{(t_l)}, S_{u_2}^{(t_l)}, R_{GS}) &= \begin{cases} (Ga, Su) & si S_{u_1}^{(t_l)}(t_l', R_q) > S_{u_2}^{(t_l)}(t_l', R_q) \\ (Su, Ga) & sinon \end{cases}
\end{aligned}$$

Nous pouvons donc en déduire la modélisation du noyau $\mathcal{K}^{R_{GS}}$ du patron de rôles suivante : $\mathcal{K}^{R_{GS}} = \{(R_q, D_p), (R_q, S_q)\}$.

2. Les rôles de Prospecteur –*Prospector*– et Mineur –*Miner*– (PM). Leur définition permet d'identifier une différence principale basée sur la nature de la tâche des rôles. Le prospecteur est plus facilement dévoué à reformuler des requêtes dans l'objectif de favoriser la diversité des résultats tandis que le Mineur est concentré sur la détection des documents pertinents. Par conséquent, nous avons identifié que le Prospecteur peut avoir tendance à être caractérisé par un taux de recouvrement R_q plus faible et un nombre de requêtes N_q plus élevé que le Mineur. Le patron de rôles $P^{R_{PM}}$ des rôles de Prospecteur, noté Pr , et de Mineur, noté Mi , est défini ainsi :

$$\begin{aligned}
P^{R_{PM}} &= \{F^{R_{PM}}, Role(S_{u_1}^{(t_l)}, S_{u_2}^{(t_l)}, R_{PM})\} \quad (5.7) \\
avec F^{R_{PM}} &= \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \\
et Role(S_{u_1}^{(t_l)}, S_{u_2}^{(t_l)}, R_{PM}) &= \begin{cases} (Pr, Mi) & si S_{u_1}^{(t_l)}(t_l', N_q) > S_{u_2}^{(t_l)}(t_l', N_q) \\ (Mi, Pr) & sinon \end{cases}
\end{aligned}$$

Nous pouvons donc en déduire le noyau $\mathcal{K}^{R_{PM}}$ du patron de rôles suivant : $\mathcal{K}^{R_{PM}} = \{(R_q, N_q)\}$.

5.4.1.5 Protocole et modèles de référence

Etant donné les objectifs de notre expérimentation, nous proposons d’analyser dans quelle mesure l’apprentissage des rôles permet d’optimiser la session de recherche en comparant l’efficacité de sessions de recherche conduites sous différents scénarios de collaboration. Plus particulièrement, nous avons adopté le protocole détaillé ci-dessous.

Nous avons divisé les sessions de RIC en intervalles de temps réguliers d’une durée d’une minute (60 secondes) dans l’objectif de détecter l’évolution des comportements des participants, leur différence et en déduire leur rôle dans la collaboration. Ainsi, pour chaque requête soumise par les participants dans un intervalle de temps $[t..t + 1]$, nous avons considéré les cinq scénarios suivants :

1. **BM25-RIC** : les documents sont ordonnés selon le modèle probabiliste BM25. Ce scénario simule une session de type médiation utilisateur dans laquelle la collaboration est réalisée par le biais de la communication uniquement et où les participants effectuent leurs recherches sur des moteurs de recherche indépendants.
2. **GS-RIC** : les documents sont ordonnés selon le modèle de RIC basé sur les rôles de Rassembleur et de Surveilleur (Shah et al., 2010). L’objectif de la comparaison avec ce scénario permet d’évaluer l’impact en terme d’efficacité d’une approche de type médiation système basée sur des rôles fixes tout au long de la session. Nous soulignons que les rôles peuvent ne pas nécessairement correspondre aux caractéristiques de comportement intrinsèques des participants.
3. **PM-RIC** : les documents sont ordonnés selon le modèle de RIC basé sur les rôles de Prospecteur et de Mineur (Pickens et al., 2008) dans l’objectif d’évaluer l’impact d’une approche de type médiation système basée sur des rôles fixes tout au long de la session. Nous soulignons que les rôles peuvent ne pas nécessairement correspondre aux caractéristiques de comportement intrinsèques des participants.
4. **Ra-RIC** : les documents sont ordonnés aléatoirement en fonction des trois modèles cités précédemment, à savoir BM25, le modèle basé sur les rôles de Rassembleur/Surveilleur (Shah et al., 2010) et celui basé sur les rôles de Prospecteur/Mineur (Pickens et al., 2008).
5. **RB-RIC** : ce scénario identifie le rôle des participants à l’instant t et applique le modèle de RIC associé. Notons que lorsqu’aucun rôle n’est détecté, le modèle BM25 est lancé.

Les quatre scénarios basés sur les rôles *GS-RIC*, *PM-RIC*, *Ra-RIC* et *RB-RIC* nécessitent un couple de requêtes, respectivement une pour chaque utilisateur. Ainsi, pour chaque requête soumise à un instant $t_l \in [t..t + 1]$ pour un utilisateur u_j , nous avons identifié la dernière requête soumise par son collaborateur $u_{j'}$.

5.4.1.6 Métriques d'évaluation

Pour chaque expérimentation utilisateurs, nous avons identifié l'ensemble des documents pertinents à partir de l'ensemble des documents distincts visités par au moins deux participants. Nous avons ainsi obtenu 159 et 47 documents pertinents respectivement pour les expérimentations US1 et US2.

Similairement aux expérimentations que nous avons menées précédemment (Chapitre 4), nous avons utilisé des mesures orientées collaboration (Shah and González-Ibáñez, 2011b) basées sur la précision et le recouvrement qui estiment l'efficacité au niveau de la session. Cependant, nous les avons modifiées afin de les rendre plus consistantes à un contexte de collaboration en combinant les mesures de micro-précision et de couverture. L'objectif est d'agréger (a) l'aspect itératif des sessions et (b) la nécessité de diversifier les ordonnancements. D'une façon plus similaire aux mesures présentées dans les travaux de Shah and González-Ibáñez (2011b), nous avons considéré des mesures de précision, rappel et F-mesure appliquées sur les ordonnancements. De plus, nous proposons d'estimer ces mesures selon deux applications du critère de diversité :

1. Critère de diversité au niveau de la session afin de vérifier que les ordonnancements retournés sur l'ensemble de la session proposent des documents distincts et pertinents :

- (a) La précision basée sur la session $Prec^S @ R(g)$ au rang R d'un groupe de collaboration g durant la session S :

$$Prec^S @ R(g) = \frac{CouvPert(L_S)}{Couverture(L_S)} \quad (5.8)$$

où $CouvPert(L_S)$ correspond au nombre de documents distincts et pertinents dans l'ensemble des listes L_S présentées lors de la session S et $Couverture(L_S)$ correspond au nombre de documents distincts dans l'ensemble des listes L_S présentées lors de la session S .

- (b) Le rappel basé sur la session $Rappel^S @ R(g)$ au rang R du groupe g durant la session S :

$$Rappel^S @ R(g) = \frac{CouvPert(L_S)}{|Pert|} \quad (5.9)$$

où $|Pert|$ exprime le nombre total de documents pertinents pour la tâche exploratoire.

- (c) La F-mesure basée sur la session $F_{syn}^S @ R(g)$ au rang R d'un groupe g basée sur la session S :

$$F^S @ R(g) = \frac{2 * Prec^S @ R(g) * Rappel^S @ R(g)}{Prec^S @ R(g) + Rappel^S @ R(g)} \quad (5.10)$$

2. Critère de diversité au niveau des itérations afin de vérifier que les ordonnancements affichés simultanément aux collaborateurs sont distincts et pertinents :

- (a) La précision basée sur les itérations $Prec^i @ R(g)$ au rang R d'un groupe de collaboration g :

$$\begin{aligned} Prec^i @ R(g) &= \frac{1}{T(g)} \sum_{t=1}^{|T(g)|} Prec^i @ R(g)^{(t_l)} \\ &= \frac{1}{T(g)} \sum_{t=1}^{|T(g)|} \frac{DistPert @ R(g)^{(t_l)}}{Dist @ R(g)^{(t_l)}} \end{aligned} \quad (5.11)$$

où $T(g)$ représente l'ensemble des instants pour lesquels des listes de documents ont été retournées. $Dist @ R(g)^{(t_l)}$, respectivement $DistPert @ R(g)^{(t_l)}$, d'un groupe g correspond au nombre total de documents distincts parmi les R premiers retournés, respectivement le nombre de documents distincts et pertinents parmi les R premiers retournés, à l'instant t pour l'ensemble des membres du groupe g .

- (b) Le rappel basé sur les itérations $Rappel^i @ R(g)$ au rang R du groupe g :

$$\begin{aligned} Rappel^i @ R(g) &= \frac{1}{T(g)} \sum_{t=1}^{|T(g)|} Recall @ R(g)^{(t_l)} \\ &= \frac{1}{T(g)} \sum_{t=1}^{|T(g)|} \frac{DistPert @ R(g)^{(t_l)}}{|Pert|} \end{aligned} \quad (5.12)$$

- (c) La F-mesure basée sur les itérations $F_{syn}^i @ R(g)$ au rang R d'un groupe g :

$$F^i @ R(g) = \frac{1}{T(g)} \sum_{t=1}^{|T(g)|} \frac{2 * Prec^i @ R(g)^{(t_l)} * Rappel^i @ R(g)^{(t_l)}}{Prec^i @ R(g)^{(t_l)} + Rappel^i @ R(g)^{(t_l)}} \quad (5.13)$$

5.4.2 Résultats

Dans cette section, nous présentons et discutons les résultats obtenus. Notre analyse est concentrée autour de trois axes principaux :

1. L'estimation de la différence entre les participants d'un même groupe de collaboration afin d'identifier si un apprentissage des rôles est possible (section 5.4.2.1).
2. Une analyse des rôles extraits grâce à notre modèle d'apprentissage des rôles (section 5.4.2.2).
3. Une évaluation de l'impact de notre modèle d'apprentissage des rôles sur l'efficacité des sessions de RIC (section 5.4.2.3).

5.4.2.1 Analyse des différences entre collaborateurs

Compte tenu de la première étape de notre modèle basée sur l'analyse des différences des comportements entre collaborateurs, nous proposons d'identifier l'évolution des comportements des participants, et plus particulièrement leurs différences au sein d'un même groupe de collaboration vis-à-vis des caractéristiques de comportement décrites dans le Tableau 5.5.

Le Tableau 5.6 illustre les moyennes des différences absolues observées entre les collaborateurs de chacun des groupes de collaboration pour chaque intervalle de temps compris entre 1 et 30 minutes après le début de la session. Ces différences sont estimées en tenant compte des caractéristiques de comportement impliquées dans les patrons de rôles décrits dans la section 5.4.1.4, à savoir le succès de la requête (S_q), le nombre de requêtes (N_q), le recouvrement entre requêtes successives (R_q) et la durée de visite (D_p).

Intervalle	1	2	3	4	5	10	15	20	25	30	
US1	S_q $ Diff $	0.535	0.746	0.801	0.837	0.820	1.004	1.023	1.101	1.240	-
	p -value		**	**	***	***	***	****	****	****	-
	N_q $ Diff $	1.153	1.339	1.746	1.932	2.339	2.898	3.508	3.983	4.190	-
	p -value		*	**	**	***	***	***	***	***	-
	R_q $ Diff $	0.309	0.340	0.413	0.421	0.380	0.291	0.303	0.259	0.238	-
	p -value		*	*	**	***	***	****	****	****	-
	D_q $ Diff $	52.725	67.169	76.289	87.344	82.109	80.544	72.196	56.760	52.567	-
	p -value		**	***	***	***	****	****	****	****	-
	S_q $ Diff $	0.333	0.467	0.548	0.699	0.661	0.811	0.986	0.570	0.400	0.107
	p -value		**	**	***	***	***	***	**	**	***
US2	N_q $ Diff $	1.200	2.200	2.100	2.500	2.700	3.100	5.800	6.500	7.000	0.500
	p -value		*	**	**	**	***	***	****	***	***
	R_q $ Diff $	0.276	0.380	0.355	0.316	0.387	0.231	0.224	0.157	0.134	0.144
	p -value		*	**	**	**	****	****	****	****	****
	D_p $ Diff $	36.558	34.285	28.116	23.635	17.386	13.435	13.988	13.062	12.554	12.779
p -value		***	***	***	***	***	***	***	***	***	

TABLE 5.6: Différences de comportement entre participants. $|Diff|$: différence absolue. p -value ($p < 0.05$) : nombre de groupes avec des différences significatives entre les participants : * : au moins un groupe. ** : au moins 50% des groupes. *** : au moins 75% des groupes. **** : tous les groupes

A partir du tableau, nous pouvons observer que, pour l'ensemble des caractéristiques de comportement, les différences entre les collaborateurs sont significatives pour au moins un groupe 1 minute après le début de la session. Ce constat nous permet de confirmer notre hypothèse que les différences entre les collaborateurs sont perceptibles assez rapidement dans la session, appuyant le fondement de notre modèle d'apprentissage des rôles. Cette conclusion est renforcée par le fait qu'entre 75% et 100% des groupes de collaboration sont caractérisés par des différences significatives 5 minutes après le début de la session.

5.4.2.2 Analyse des changements de rôles détectés au cours de la session

Dans cette section, nous étudions comment notre modèle d'apprentissage des rôles agit dans une session de recherche collaborative. Dans cet objectif, nous définissons trois mesures, présentées dans l'Illustration 5.6, dont l'objectif est d'analyser le changement des rôles détectés au cours de la session :

1. Le **nombre de couples de rôles**, noté $N_{couples}$, normalisé par le nombre d'intervalles, respectivement 25 et 30 pour US1 et US2. Cette métrique analyse la stabilité des différences de comportement entre les participants d'un même groupe tout au long d'une session de recherche. Plus particulièrement, elle mesure le nombre de changements des couples de rôles successivement détectés par notre modèle, auquel nous ajoutons 1 pour le rôle neutre, appelé SansRole, considéré au début de la session. Nous faisons l'hypothèse qu'il y a un changement entre l'instant t et l'instant $t + 1$ lorsque les paires de rôles assignées pour ces deux instants au groupe de collaborateurs sont différentes. Une valeur proche de 1 signifie que les paires de rôles détectées sont différentes à chaque instant tandis qu'une valeur proche de 0 traduit un comportement constant tout au long de la session de la paire de collaborateurs. Plus formellement, le nombre $N_{couples}$ de couples de rôles est estimé ainsi :

$$N_{couples} = \frac{1 + \sum_{t=1}^{T-1} \mathbb{1}_{\{\mathcal{R}(t) \neq \mathcal{R}(t+1)\}}}{T} \quad (5.14)$$

où T représente le nombre d'instants dans la session S séparés par des intervalles réguliers de 1 minute et $\mathcal{R}(t)$ exprime la paire de rôles détectée à l'instant t de la session S .

2. Le **nombre de rôles distincts**, noté $N_{rôles}$, assigné à un utilisateur tout au long de la session, normalisé par le nombre d'intervalles. Nous considérons cinq possibilités de rôles : Prospecteur, Mineur, Rassembleur, Surveilleur et SansRole. Ce dernier correspond au cas où aucun rôle n'a été détecté. Nous soulignons que la valeur maximale qu'un utilisateur peut obtenir s'il reçoit chacun de ces cinq rôles est respectivement égale à 0.2 pour US1 et 0.16 pour US2. Le nombre $N_{rôles}$ de rôles distincts est estimé de la façon suivante :

$$N_{rôles} = \frac{\mathcal{R}(u)}{T} \quad (5.15)$$

où $\mathcal{R}(u)$ représente le nombre de rôles distincts assignés à l'utilisateur u . Notons que de façon indifférenciée, le nombre de rôles de chacun des utilisateurs est équivalent.

3. Le **ratio couple-rôle**, noté R_{change} , permet d'analyser si, lorsqu'une paire de rôles est détectée plusieurs fois pour un groupe, la répartition des rôles à la paire de collaborateurs est stable ou si les participants interchangent leur

comportement de façon à recevoir des rôles différents au sein de cette paire de rôles. Par exemple, un utilisateur peut être détecté comme Prospecteur en début de session et comme Mineur en fin de session. Dans cet objectif, nous établissons le ratio R_{change} entre le nombre de couples de rôles $N_{couples}$ et le nombre de rôles distincts $N_{rôles}$ pour chacune des sessions de recherche collaborative, comme présenté ci-dessous :

$$R_{change} = \frac{N_{couples}}{N_{rôles}} \quad (5.16)$$

Une valeur égale à 1 met en évidence le fait que, même si les paires de rôles détectées tout au long de la session varient et sont redondantes, les rôles à l'intérieur de chacune des paires de rôles sont constamment assignés au même participant. En d'autres termes, si un participant est détecté Rassembleur à l'instant t de la session et que la paire de rôles Rassembleur-Surveilleur est détectée à l'instant $t + 4$, le rôle de Rassembleur est une nouvelle fois détecté pour ce même participant. A l'inverse, une valeur inférieure à 1 traduit le fait que pour au moins une paire de rôles détectée à deux instants différents, chaque participant a successivement été assigné aux différents rôles associés à la paire de rôles. Par exemple, un utilisateur peut recevoir le rôle de Rassembleur à l'instant t et de Surveilleur à l'instant $t + 4$. Par contre, un ratio supérieur à 1 signifie que les paires de rôles détectées dans la session varient plus que la distribution des rôles à chaque participant au sein de cette même paire, lorsqu'elle est détectée à plusieurs reprises dans la session.

Illustration 5.6 Analyse des rôles des collaborateurs.

Nous considérons une session de recherche collaborative S de 6 minutes impliquant deux collaborateurs. Notre modèle d'apprentissage a été appliqué et nous avons pu détecter les rôles des collaborateurs pour chaque minute de la session, comme présenté dans le tableau ci-dessous :

instant	1	2	3	4	5	6
rôles	SansRole	GS	GS	PM	GS	PM
rôle u_1	SansRole	gatherer	gatherer	mineur	surveyor	mineur
rôle u_2	SansRole	surveyor	surveyor	prospector	gatherer	prospector

A partir de l'analyse des rôles, nous obtenons les valeurs suivantes :

- $N_{couples} = 5/6 = 0.83$
- $N_{rôles} = 4/6 = 0.66$
- $R_{change} = 0.83/0.66 = 1.25$

Nous pouvons conclure que les paires de rôles ainsi que les rôles attribués au sein de la session varient avec une prédominance pour la variation des paires de rôles plutôt que les rôles intrinsèques aux utilisateurs.

Le Tableau 5.7 présente les statistiques de chacune de ces trois mesures pour chacune des expérimentations utilisateurs. Plus particulièrement, nous avons estimé la valeur moyenne et l'intervalle de valeur, sous le format “moyenne [$min - max$]”.

	US1	US2
$N_{couples}$	0.13 [0.04 – 0.28]	0.08[0.03 – 0.2]
$N_{rôles}$	0.09[0.04 – 0.16]	0.09[0.04 – 0.16]
R_{change}	0.88[0.29 – 1.5]	1.18[0.5 – 1.5]

TABLE 5.7: Analyse de l'évolution des rôles détectés lors des sessions de recherche

Pour les deux expérimentations, nous pouvons observer, dans le Tableau 5.7, (a) le faible nombre de changements en terme de couples $N_{couples}$ et de rôles distincts $N_{rôles}$ et (b) une valeur proche de 1 pour le ratio couple-rôle R_{change} mettant en évidence une certaine stabilité dans les paires de rôles assignées à la paire de collaborateurs ainsi que dans les rôles assignés aux participants au sein de la même paire de rôles. En effet, le nombre normalisé $N_{couples}$ de couples, respectivement 0.13 et 0.08 pour les expérimentations US1 et US2, reflète que les comportements des utilisateurs sont généralement stables tout au long de la session. Nous soulignons le fait que le comportement des participants de la première expérimentation US1 est moins stable que celui des participants de l'expérimentation US2. Les deux expérimentations ont des résultats équivalents en terme de nombre de rôles distincts détectés pour chaque participant avec une valeur égale à 0.09. Par conséquent, nous obtenons des conclusions sensiblement opposées avec le ratio couple-rôle, bien que les valeurs soient pour les deux expérimentations proches de 1. Il semble donc que les paires de participants de l'expérimentation US1 sont sensiblement plus variables en terme de comportement global qu'en terme de comportement individuel. En effet, nous avons détecté plus de changement dans les paires de rôles assignées au groupe de collaborateurs que dans l'assignation des rôles pour un participant unique. Par contre, pour l'expérimentation US2, le comportement des participants est plus fluctuant que le comportement global de la paire de collaborateurs dans la mesure où, pour une paire de rôles détectée à deux instants différents, le même participant a été assigné à deux rôles différents.

Cette analyse nous conduit à explorer si la fenêtre de temps d'apprentissage des rôles impacte l'efficacité des sessions de RIC. Nous avons ainsi estimé l'efficacité de notre scénario, noté *RB-RIC*, qui repose sur notre modèle d'apprentissage des rôles afin d'exécuter le modèle d'ordonnancement collaboratif adapté au comportement des utilisateurs, en réalisant l'apprentissage pour des intervalles variant entre 1 et 5 minutes. Le Tableau 5.8 illustre les résultats obtenus selon la F-mesure au niveau des itérations afin d'analyser plus précisément l'impact de la distribution des rôles à un moment donné. Avant toute analyse, nous soulignons le fait que la métrique d'évaluation est faible compte tenu du faible nombre de documents pertinents identifiés pour chaque expérimentation. Nous pouvons observer dans le Tableau 5.8 que pour l'expérimentation US2, même si l'efficacité semble être plus élevée pour la fenêtre temporelle de 5 minutes, les résultats ne sont pas significativement différents, compte tenu du faible nombre de groupes.

Cependant, pour la première expérimentation, la valeur optimale correspond à la fenêtre de temps d'une minute. De plus, les tests de significativité, réalisés par le test de Student, montrent que ce résultat est significativement différent des autres configurations reposant sur des fenêtres de 2, 3, 4 et 5 minutes, d'autant plus qu'ils ne sont pas significativement différents entre eux. Ce résultat met en évidence le fait que même si les rôles ne varient pas significativement, il reste cependant plus efficace de proposer un apprentissage des rôles pour chaque minute dans l'objectif de mieux s'adapter et de proposer des modèles d'ordonnancement qui correspondent mieux au comportement des utilisateurs, permettant ainsi d'optimiser la session de recherche. Pour cette raison, nous considérons dans le reste de nos expérimentations une fenêtre d'apprentissage d'une minute.

	1	2	3	4	5
US1	0.01673	0.01650	0.01654	0.01649	0.01647
US2	0.07061	0.06683	0.06931	0.06650	0.07203

TABLE 5.8: Impact de la fenêtre temporelle d'apprentissage sur l'efficacité des sessions

En complément à ces analyses, nous avons également effectué une analyse des corrélations entre le ratio couple-rôle et la F-mesure. Les résultats ont montré des corrélations non significatives entre les deux variables avec des valeurs respectivement de 0.003 et -0.003 pour les expérimentations US1 et US2. Ce résultat suggère que l'efficacité des sessions repose plus sur la qualité d'apprentissage que sur la quantité de rôles distincts détectés.

5.4.2.3 Analyse de l'impact de l'apprentissage des rôles sur l'efficacité au niveau de la session

Dans cette section, nous mesurons et discutons de l'impact de notre modèle d'apprentissage des rôles sur l'efficacité globale des sessions de RIC. Notons que nous fixons $\theta = 0.05$ comme valeur de significativité au test de Kolmogorov-Smirnov utilisé dans l'étape de différenciation des participants.

Dans un premier temps, nous avons estimé les mesures d'efficacité en fonction du critère de diversité calculé au niveau de la session (Equations 5.8, 5.9 et 5.10). Nous avons considéré les différents scénarios, à savoir notre modèle d'apprentissage des rôles, noté *RB-RIC*, et les quatre modèles de références : *BM25-RIC*, *GS-RIC*, *PM-RIC* et *Ra-RIC*, décrits dans la section 5.4.1.5. Les résultats sont présentés dans le Tableau 5.9 et montrent l'efficacité de notre modèle qui fournit des améliorations significatives comprises entre 10.245% et 90.810% sur l'ensemble des scénarios. Par contre, les améliorations notées pour l'expérimentation US2 sont moins significatives du fait du nombre de groupes de collaboration dont la conséquence est qu'atteindre la significativité des tests de student est plus difficile. Cependant, nous observons que notre système semble

plus efficace pour une réponse à une tâche orientée rappel à l'inverse des modèles basés sur la collaboration *GS-RIC*, *PM-RIC* et *Ra-RIC*. Nous pouvons également noter que notre modèle n'est pas significativement plus efficace qu'un modèle basé sur une médiation utilisateur seulement, *BM25-RIC*. Une analyse des mesures au niveau itérations complète ce constat.

	$Prec^S@20$			$Rappel^S20$			$F^S@20$		
	value	%Cg	p-value	value	%Tx	p-value	value	%Tx	p-value
BM25-RIC	0.036	3.804		0.010	0		0.015	13.33	
GS-RIC	0.039	-2.564		0.0006	41.200	***	0.012	27.65	***
US1 PM-RIC	0.048	-20.833		0.0.005	90.810	***	0.011	32.91	***
Ra-RIC	0.038	0		0.008	13.141	*	0.014	10.245	*
RB-RIC	0.038	-		0.010	-		0.017	-	
BM25-RIC	0.057	4.792		0.017	0		0.026	0	
GS-RIC	0.050	20.032		0.009	70	*	0.015	63.309	*
US2 PM-RIC	0.089	-32.80		0.012	30.769		0.021	17.489	
Ra-RIC	0.053	11.994		0.014	17.241		0.022	16.718	
RB-RIC	0.060	-		0.017	-		0.026	-	

TABLE 5.9: Comparaison de l'impact de notre algorithme d'apprentissage des rôles sur l'efficacité d'une tâche de RIC en fonction des mesures orientées session. %Tx : taux d'amélioration de *RB-RIC*. Significativité du test de Student * : $0.01 < t \leq 0.05$; ** : $0.001 < t \leq 0.01$; *** : $t \leq 0.001$

A cet effet, le Tableau 5.10 présente une analyse comparative des résultats obtenus lorsque le principe de diversité est analysé au niveau des itérations (Equations 5.12, 5.13 et 5.13). Les résultats mettent également en évidence que notre modèle fournit des augmentations significatives par rapport à la plupart des mesures basées sur la collaboration, comme constaté dans le Tableau 5.9. Pour l'expérimentation US1, l'efficacité d'une session collaborative est améliorée grâce à notre modèle avec des taux compris entre +10.408% et +18.316% pour la mesure de précision, à l'exception d'une diminution non significative par rapport au scénario basé sur la paire de rôles Prospecteur-Mineur. Cette différence entre mesures orientées session et itération souligne la capacité du modèle *PM-RIC* à fournir simultanément des ordonnancements incluant des documents pertinents aux utilisateurs et respectant les principes de la collaboration. Nous notons dans cette analyse que notre modèle fournit des résultats pour l'expérimentation US1 qui sont significativement supérieurs au modèle de référence *BM25*, à l'inverse de l'analyse des mesures orientées sessions. Cela signifie que notre modèle, du fait de la prise en compte de la complémentarité des compétences et du principe de division du travail fournit des ordonnancements, lorsqu'ils sont simultanément affichés aux différents collaborateurs, à la fois distincts et pertinents. Pour l'expérimentation US2 qui présente un plus petit nombre de groupes de collaborations, nous avons obtenu des améliorations significatives comprises entre 52.786% et 63.818% par rapport au modèle de référence *GS-RIC* basé sur les rôles de Rassembleur et de Surveilleur.

		<i>Prec</i> ⁱ @20			<i>Rappel</i> ⁱ @20			<i>F</i> ⁱ @20	
		value	% <i>Cg</i>	<i>p</i> -value	value	% <i>Tx</i>	<i>p</i> -value	value	% <i>Tx</i> <i>p</i> -value
US1	BM25-RIC	0.041	10.408	*	0.010	4.636	*	0.016	5.372
	GS-RIC	0.038	18.316	***	0.008	25.504	***	0.014	24.521 ***
	PM-RIC	0.050	-9.482		0.012	-13.991		0.019	-13.397
	Ra-RIC	0.041	11.484	*	0.009	12.895	*	0.015	12.777 *
	RB-RIC	0.045	-		0.010	-		0.017	-
US2	BM25-RIC	0.075	3.347		0.063	2.586		0.069	2.833
	GS-RIC	0.058	34.636		0.040	63.818	*	0.046	52.786 *
	PM-RIC	0.092	-16.051		0.078	-16.493		0.084	-16.317
	Ra-RIC	0.070	10.714		0.056	16.201		0.062	14.324
	RB-RIC	0.077	-		0.065	-		0.071	-

TABLE 5.10: Comparaison de l’impact de notre algorithme d’apprentissage des rôles sur l’efficacité d’une tâche de RIC en fonction des mesures orientées itération. %*Tx* : taux d’amélioration de *RB-RIC*. Significativité du test de Student * : $0.01 < t \leq 0.05$; ** : $0.001 < t \leq 0.01$; *** : $t \leq 0.001$

Plus généralement, nous notons les tendances suivantes pour les deux expérimentations :

1. Une session de recherche collaborative guidée par l’apprentissage des rôles permet d’obtenir des résultats qui dépassent l’agrégation des résultats individuels mise en évidence par le scénario *BM25-RIC*. Cela souligne la nécessité de bénéficier des différences entre utilisateurs pour apprendre leurs rôles afin d’optimiser une session de recherche.
2. Notre scénario basé sur l’apprentissage des rôles fournit de meilleurs résultats que celui dont le modèle d’ordonnancement collaboratif basé sur les rôles de Rassembleur et de Surveilleur. Nous pouvons ainsi supposer que notre scénario est mieux adapté qu’un scénario où le modèle d’ordonnancement collaboratif repose sur des rôles de Rassembleur et de Surveilleur fixés sur l’ensemble de la session et où l’ordonnancement est réalisé par une simple fusion et classification des documents retournés par les requêtes des deux collaborateurs (Shah et al., 2010).
3. L’algorithme basé sur les rôles de Prospecteur et de Mineur semble fournir de meilleurs résultats que notre scénario lorsque les ordonnancements sont considérés par itération. Cependant, ces améliorations ne sont pas significatives. Une explication possible à cette équivalence en terme d’efficacité peut être liée au fait que le modèle de Prospecteur-Mineur analyse les documents retournés et les requêtes soumises depuis le début de la session dans l’objectif d’optimiser l’ordonnancement des documents. D’une façon similaire, notre modèle repose sur les actions des utilisateurs afin de détecter les rôles des collaborateurs dans l’objectif d’adapter dynamiquement les modèles d’ordonnancement en conséquence. Néanmoins, notre modèle reste plus efficace lorsque les ordonnancements sont considérés dans la globalité, selon les mesures de rappel et F-mesure orientées session.

4. Notre modèle semble plus efficace qu'un scénario dans lequel les rôles sont assignés de façon aléatoire. Cela met en évidence la fiabilité de notre méthodologie dont l'objectif est de détecter les rôles adaptés au comportement des collaborateurs, et ce, afin d'optimiser l'efficacité de la session.

Par conséquent, nous pouvons conclure que notre modèle d'apprentissage des rôles dans une session de RIC optimise la collaboration et améliore l'effet synergique des collaborateurs.

5.4.2.4 Analyse de l'impact de l'apprentissage des rôles sur l'efficacité au niveau de l'itération

Notre objectif ici consiste en une analyse plus en profondeur de l'efficacité de notre modèle tout au long de la session de RIC. Ainsi, nous illustrons dans la Figure 6.7 l'efficacité, estimée par la valeur moyenne de la F-mesure $F^i@20$, des cinq scénarios d'expérimentation entre chaque intervalle d'itérations $[t..t + 1]$.

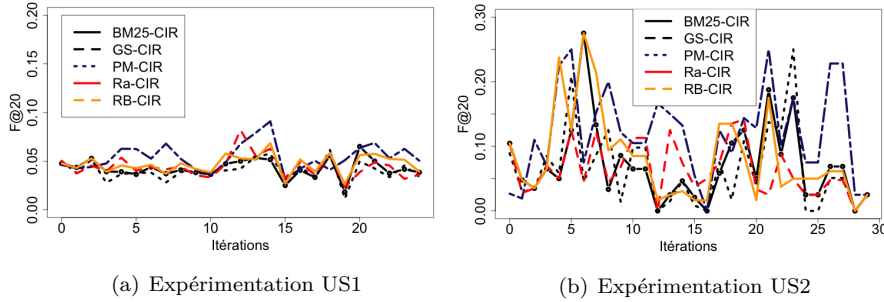


FIGURE 5.8: Comparaison de l'efficacité sur l'ensemble de la session de recherche.

Pour la première expérimentation US1, la courbe de notre approche d'apprentissage des rôles, notée *RB-RIC*, dépasse généralement les courbes des modèles de référence, à l'exception de la courbe du scénario basé sur les rôles de Prospecteur et de Mineur, comme attendu à la suite des résultats obtenus précédemment. Ce constat est moins identifiable dans la Figure 5.8(b) du fait du faible nombre de groupes qui sont impliqués dans l'expérimentation US2. Ainsi, pour cette expérimentation, les différences entre notre modèle et les scénarios de référence sont plus variables et non significatives. Néanmoins, nous pouvons conclure que notre modèle permet d'optimiser le processus de collaboration. Un aspect intéressant à noter est la tendance à la décroissance de notre scénario *RB-RIC* à la fin des sessions de recherche. Nous pouvons expliquer cela par le fait que les sessions de recherche sont assez longues, entre 25 et 30 minutes et que nous avons noté pendant les expérimentations que les participants étaient moins impliqués à la fin de la session. Nous pouvons ainsi faire l'hypothèse que l'effort cognitif pour la reformulation des requêtes était plus élevé, ce qui impacte négativement

l'expression du besoin en information et par conséquent la fiabilité des résultats issus des algorithmes d'apprentissage des rôles.

Résumé

Dans ce chapitre, nous proposons un modèle supervisé d'apprentissage de rôles basé sur les actions des collaborateurs. Nous faisons l'hypothèse que les rôles des collaborateurs peuvent varier au cours de la session en fonction des interactions avec les documents ou les autres membres du groupe. Cette hypothèse a été motivée par une analyse préliminaire réalisée sur une expérimentation utilisateurs, qui a également mis en évidence l'impact de la contrainte des rôles sur le comportement des collaborateurs. L'objectif de ce modèle est d'adapter dynamiquement les modèles d'ordonnancement aux rôles des collaborateurs afin de bénéficier des compétences pour lesquelles ils sont le plus spécialisés par rapport aux autres membres du groupe. Ainsi, la session de recherche est perçue comme une session hybride où la médiation est à la fois utilisateur et système, laissant les collaborateurs libres de leurs actions qui permettent au système de s'adapter aux caractéristiques des utilisateurs pour l'ordonnancement des documents. Plus spécifiquement, notre modèle d'apprentissage s'applique à une paire d'utilisateurs et nous aide à comprendre dans quelle mesure les collaborateurs sont différents vis-à-vis des caractéristiques de recherche. L'évaluation expérimentale menée repose sur des fichiers logs de deux expérimentations utilisateurs et a permis de démontrer le bénéfice de l'apprentissage des rôles prédéfinis dans un processus de collaboration.

Cette contribution a cependant quelques limites car elle considère une approche supervisée où un ensemble prédéfini de rôles peut être détecté. En effet, il est possible que les rôles ne correspondent pas exactement aux caractéristiques des collaborateurs et, comme indiqué dans notre analyse préliminaire, biaise négativement le comportement de l'utilisateur dans la session. Ceci peut impacter l'efficacité de la session. Pour répondre à cette problématique, nous proposons dans le Chapitre 6 une approche non supervisée qui dérive les patrons de rôles latents au cours de la session à partir des actions des collaborateurs sans fixer a priori des rôles possibles. Ces rôles latents sont ensuite réinjectés dans un modèle d'ordonnancement collaboratif qui permet de distribuer les documents aux collaborateurs en fonction de leurs compétences.

Chapitre 6

Modèle de RIC non-supervisé basé sur les méta-rôles des collaborateurs

Introduction

En complément de la contribution précédente qui présente un modèle d'apprentissage supervisé des rôles des collaborateurs, nous nous intéressons dans ce chapitre à une approche non-supervisée qui a pour objectif d'identifier les rôles latents des collaborateurs, appelés aussi méta-rôles, à partir de leurs actions dans la session de recherche afin de les injecter dans l'ordonnancement collaboratif des documents.

Le chapitre est organisé de la façon suivante. La section 6.1 présente le contexte et introduit les questions de recherche de notre contribution. La section 6.2 définit les notations utilisées par la suite. La section 6.3 définit notre modèle d'apprentissage des rôles. Enfin, l'évaluation expérimentale menée afin de prouver l'efficacité de notre modèle est détaillée dans la section 6.4.

6.1 Contexte et questions de recherche

Nous avons abordé précédemment la nécessité de considérer les sessions de recherche de façon hybride, qui combinent une médiation système et une médiation utilisateur, afin de bénéficier des compétences complémentaires des colla-

borateurs et d'éviter de les contraindre dans des objectifs de tâche qui ne leur correspondent pas entièrement. En effet, nous avons démontré lors de l'analyse préliminaire de la section 5.1.2 que les collaborateurs peuvent affiner leurs compétences et leurs actions au fur et à mesure de la session de recherche. Ainsi, les modèles d'ordonnancement collaboratifs de l'état de l'art (Pickens et al., 2008; Shah et al., 2010), reposant sur une assignation de rôles fixes sur l'ensemble de la session, ne prennent pas en compte cet aspect. Pour résoudre ce problème, nous avons proposé, dans notre précédente contribution, un modèle d'apprentissage supervisé des rôles (Soulie et al., 2014a) permettant (a) d'identifier dynamiquement le rôle des collaborateurs, parmi un ensemble de rôles prédéfinis issus de la littérature, et (b) d'ordonner les documents grâce au modèle de RIC associé. Cependant, la problématique de l'inadéquation des compétences des collaborateurs avec les rôles prédéfinis qui peuvent être détectés est une des principales limites de cette contribution. Également, d'un point de vue de l'exploitation des compétences complémentaires, un collaborateur peut correspondre à plusieurs rôles au sein d'une même session. Ainsi, assigner le meilleur rôle à un collaborateur ne permet pas de bénéficier pleinement de ses compétences. Dans ce contexte, nous proposons de généraliser ce précédent modèle selon une approche non supervisée, appelée MineRank, qui identifie de façon dynamique les rôles latents des collaborateurs à partir des actions de la paire de collaborateurs afin d'améliorer l'efficacité des ordonnancements. Nous soulignons le fait que ces rôles ne sont pas nécessairement définis dans une taxonomie (Golovchinsky et al., 2009b; Pickens et al., 2008; Shah et al., 2010).

Notre contribution est guidée par les questions de recherche suivantes :

1. *Comment les actions des utilisateurs permettent-elles de caractériser les méta-rôles ?* Similairement à notre contribution précédente, les rôles latents sont identifiés à partir des différences de comportement entre collaborateurs, et plus particulièrement une analyse des corrélations des différences. La spécificité de cette contribution étant de considérer des rôles non prédéfinis, l'objectif principal est de faire émerger les rôles latents des collaborateurs qui permettent d'optimiser la complémentarité de leurs actions sans dégrader l'efficacité de la session. Par conséquent, nous proposons une méthode de sélection de variables adaptée au contexte collaboratif permettant d'identifier les caractéristiques de comportement discriminantes qui maximisent la différence entre les utilisateurs et leurs préférences vis-à-vis de la tâche de recherche.
2. *Comment injecter les méta-rôles dans un modèle d'ordonnancement afin d'augmenter l'efficacité de la session ?* A ce niveau, une des difficultés repose sur le fait qu'il n'existe aucun modèle d'ordonnancement basé sur des rôles latents, non prédéfinis, pouvant intégrer les caractéristiques liées à la complémentarité des compétences des collaborateurs et à la division du travail. Ainsi, nous proposons d'utiliser un algorithme d'apprentissage qui, à partir des méta-rôles des collaborateurs, prédit la pertinence des documents vis-à-vis des collaborateurs afin d'identifier quel utilisateur est le plus à même de les juger.

6.2 Notations : modélisation des méta-rôles

Nous présentons les principaux concepts sous-jacents de notre contribution. Plus particulièrement, nous définissons la session de recherche ainsi que la notion de méta-rôles des collaborateurs. Compte tenu du contexte de généralisation de la contribution présentée dans le Chapitre 5, certaines notations considérées dans ce chapitre sont extraites des notations détaillées dans la section 5.2.

Session de recherche. Nous considérons une session de recherche collaborative synchrone S où une paire d'utilisateurs u_1 et u_2 interagissent afin de résoudre un besoin en information partagé I pendant une durée T . Chaque utilisateur $u_j \in \{u_1, u_2\}$ effectue des recherches séparément et formule ses propres requêtes $\{q_{u_j}^1, q_{u_j}^2, \dots, q_{u_j}^T\}$, ordonnées de façon chronologique, dans l'objectif d'identifier des documents pertinents dans les ensembles de résultats respectifs $\{D_{u_j}^1, D_{u_j}^2, \dots, D_{u_j}^T\}$. Nous avons formulé dans le chapitre précédent l'hypothèse de modéliser les actions des utilisateurs par un ensemble $F = \{f_1, \dots, f_k, \dots, f_n\}$ de caractéristiques de comportement f_k . Compte tenu de l'objectif de notre contribution d'ordonner les documents en fonction du comportement des collaborateurs, nous avons choisi des caractéristiques basées sur le document extraites de précédents travaux (Agichtein et al., 2006; White and Dumais, 2009) et détaillées dans le Tableau 6.1.

	Description	Mesure
	R_{Ti} Recouvrement Titre-Requête	Ratio moyen de mots communs entre le titre de la page et la requête
\mathcal{C}^R	R_{Te} Recouvrement Texte-Requête	Ratio moyen de mots communs entre le texte de la page et la requête
	R_{An} Recouvrement Annotation-Requête	Ratio moyen de mots communs entre l'annotation de la page et la requête
	R_{Sn} Recouvrement Snippet-Requête	Ratio moyen de mots communs entre le snippet de la page et la requête
	RV_R Rang de visite	Rang de visite de la page par rapport à la requête
	D_{PR} Durée page-Requête	Temps en secondes entre la soumission de la requête et la visite de la page
\mathcal{C}^R	D_p Durée de visite	Durée moyenne en secondes d'une visite sur une page
	D_{Do} Durée de visite du domaine	Durée cumulée du temps de visite en secondes d'un domaine
	Lis Lisibilité (Kincaid et al., 1975)	Lisibilité du contenu du document
	$Spec$ Spécificité (Kim, 2006)	Spécificité du contenu du document
	$Note$ Note	Note de la page

TABLE 6.1: Caractéristiques de comportement

Ces caractéristiques sont capturées durant la session de recherche et sont utilisées pour construire la matrice temporelle $S_{u_j}^{(t_l)} \in \mathcal{R}^{t_l \times n}$ du comportement de l'utilisateur u_j à un instant particulier t_l de la session de recherche. Chaque élément $S_{u_j}^{(t_l)}(t_{l'}, f_k)$ de la matrice $S_{u_j}^{(t_l)}$ représente la valeur moyenne de la caractéristique f_k pour l'utilisateur u agrégée pour l'intervalle de temps $[0..t_{l'}]$ compte tenu des documents sélectionnés $\mathcal{D}(u_j)^{(t_{l'})}$ durant cette période. Similairement aux hypothèses formulées dans le Chapitre 5, la modélisation temporelle du comportement des utilisateurs $S_{u_j}^{(t_l)}$ permet de prendre en compte le fait que leur comportement s'affine au cours de la session et de les caractériser d'une façon globale à un instant t_l en réduisant le biais d'actions bruitées dans la session.

Méta-rôles des collaborateurs. Les méta-rôles des collaborateurs correspondent aux rôles latents des collaborateurs à un instant donné de la session. Ils sont caractérisés par des patrons de méta-rôles qui modélisent le comportement des utilisateurs. Similairement au Chapitre 5, nous formulons les hypothèses suivantes liées aux patrons de méta-rôles :

1. Les patrons de méta-rôles modélisent les complémentarités entre collaborateurs vis-à-vis de leur comportement ou leurs compétences de recherche au cours d'une session.
2. Les complémentarités et les similarités de comportement sont traduites par des corrélations (négatives ou positives) entre les caractéristiques de comportement.
3. La considération des différences entre collaborateurs vis-à-vis de chacune des caractéristiques dans le calcul des corrélations permet d'identifier les compétences pour lesquelles chaque collaborateur est le plus efficace.

Dans ce chapitre, nous avons introduit le besoin de modéliser des rôles implicites, non nécessairement labélisés. Ainsi, dans la suite de notre contribution, nous considérons des patrons de méta-rôles latents, sans aucun a priori de rôles et non nécessairement identifiés dans une taxonomie (Golovchinsky et al., 2009b), appelés également méta-rôles. Nous définissons le patron du méta-rôle $MR_{1,2}^{(t_l)}$ entre deux rôles latents (r_1, r_2) qui modélise le comportement de deux utilisateurs u_1 et u_2 à un instant t_l selon un ensemble de deux éléments :

1. Un noyau de caractéristiques de comportement $\mathcal{K}_{1,2}^{(t_l)} = \{f_1, f_2, \dots, f_p\} \in F^p$ qui recense pour la paire de collaborateur (u_1, u_2) les caractéristiques de comportement impliquées dans le patron de rôles.
2. Une matrice de corrélations $C_{1,2}^{(t_l)}$ qui met en évidence les complémentarités et similarités de comportement entre les deux utilisateurs u_1 et u_2 . Les corrélations sont estimées sur la base de différences $\Delta_{1,2}^{(t_l)}(f_k)$ et $\Delta_{1,2}^{(t_l)}(f_{k'})$ des deux caractéristiques de comportement f_k et $f_{k'}$ entre les deux collaborateurs u_1 et u_2 , où $\Delta_{1,2}^{(t_l)}(f_k) = S_{u_1}^{(t_l)}(f_k) - S_{u_2}^{(t_l)}(f_k)$. Notons que la corrélation entre une paire de caractéristiques est effectuée seulement si ces deux caractéristiques ont été détectées comme significativement différentes vis-à-vis des

deux collaborateurs grâce au test de Kolmogorov-Smirnov. Chaque élément de la matrice $C_{1,2}^{(t_i)}(f_k, f_{k'})$ est estimé comme suit :

$$C_{1,2}^{(t_i)}(f_k, f_{k'}) = \begin{cases} \rho(\Delta_{1,2}^{(t_i)}(f_k), \Delta_{1,2}^{(t_i)}(f_{k'})) & \text{si } p_{u_1, u_2}^{(t_i)}(f_k) < \theta \\ & \text{et } p_{u_1, u_2}^{(t_i)}(f_{k'}) < \theta \\ 0 & \text{sinon} \end{cases} \quad (6.1)$$

où $p_{u_1, u_2}^{(t_i)}(f_k)$ correspond à la valeur de significativité du test de Kolmogorov-Smirnov associé à la caractéristique f_k et θ représente le seuil de significativité associé au test. $\rho(\Delta_{1,2}^{(t_i)}(f_k), \Delta_{1,2}^{(t_i)}(f_{k'}))$ exprime la corrélation entre les deux vecteurs de différences $\Delta_{1,2}^{(t_i)}(f_k)$ et $\Delta_{1,2}^{(t_i)}(f_{k'})$.

Compte tenu de notre objectif de faire émerger les complémentarités entre collaborateurs, nous faisons l'hypothèse que deux caractéristiques f_k et $f_{k'}$ sont redondantes si la corrélation $C_{1,2}^{(t_i)}(k, k')$ de leur différence est proche de 1. A l'inverse, plus la valeur de corrélation est proche de -1, plus les collaborateurs sont complémentaires vis-à-vis de ces deux caractéristiques de comportement.

6.3 Modèle non supervisé de collaboration basé sur les méta-rôles

L'objectif de notre modèle repose sur l'hypothèse d'une collaboration basée sur une session de recherche hybride où la médiation est à la fois de type utilisateur et système permettant de guider les collaborateurs tout en les laissant libres d'exprimer leur complémentarité en terme de compétences de recherche. Pour cela, nous proposons une méthodologie en deux étapes, comme illustré dans la Figure 6.2 :

1. Un apprentissage des méta-rôles qui étudie le comportement des collaborateurs à partir de caractéristiques de comportement dans l'objectif d'identifier celles pour lesquelles les collaborateurs sont complémentaires. Cette étape consiste en une sélection des caractéristiques de comportement les plus discriminantes, représentatives du méta-rôle, et qui maximisent les différences entre les utilisateurs.
2. Un modèle d'ordonnancement collaboratif qui réinjecte les méta-rôles des collaborateurs afin d'optimiser la session de recherche. Pour cela, nous proposons de considérer les caractéristiques de comportement issues du noyau des méta-rôles comme sources d'évidence dans l'assignation des documents aux collaborateurs.

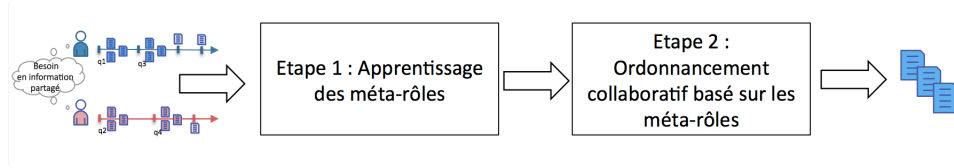


FIGURE 6.1: Etapes du modèle collaboratif non supervisé basé sur les méta-rôles des collaborateurs

6.3.1 Apprentissage des méta-rôles dans une session de RIC

L'objectif des méta-rôles est d'identifier les comportements de recherche qui sont les plus complémentaires pour être ensuite exploités dans un modèle d'ordonnancement collaboratif. Comme énoncé précédemment, la complémentarité peut être traduite par la maximisation des différences entre les utilisateurs en termes de préférences des documents et de comportement de recherche sans pour autant dégrader la qualité d'ordonnancement que ces caractéristiques peuvent fournir si elles étaient utilisées comme sources d'évidence d'un modèle de recherche d'information.

Dans ce contexte, nous proposons une approche d'apprentissage des méta-rôles des collaborateurs basée sur un algorithme de sélection des caractéristiques adapté au contexte de RIC. La problématique sous-jacente de cette approche consiste en l'identification des caractéristiques les plus discriminantes qui caractérisent le comportement des collaborateurs et leur complémentarité au sein de la session. L'intuition de notre méthode d'apprentissage des méta-rôles est illustrée dans la Figure 6.2. La sélection des caractéristiques repose sur une analyse des comportements des utilisateurs. Similairement aux notions abordées dans le Chapitre 5, les comportements sont comparés entre membres d'un même groupe de collaboration dans l'objectif de mettre en évidence les différences des comportements entre collaborateurs. Une fois que les utilisateurs sont identifiés comme différents vis-à-vis des caractéristiques de comportement, nous proposons de modéliser un réseau pondéré de ces caractéristiques de comportement. Ce dernier traduit la complémentarité des collaborateurs vis-à-vis d'une paire de caractéristiques et la qualité d'une caractéristique à être utilisée comme source d'évidence dans un modèle collaboratif. L'objectif principal est d'identifier grâce à un algorithme basé sur la théorie des graphes celles qui sont les plus importantes et discriminantes pour caractériser les méta-rôles des collaborateurs en temps réel dans une session de RIC.

Dans ce qui suit, nous présentons le cadre de sélection des caractéristiques basé sur un problème d'optimisation ainsi que ces hypothèses sous-jacentes. Ce dernier repose sur des mesures de qualité détaillées par la suite. Enfin, nous présentons une solution au problème d'optimisation basée sur la théorie des graphes.

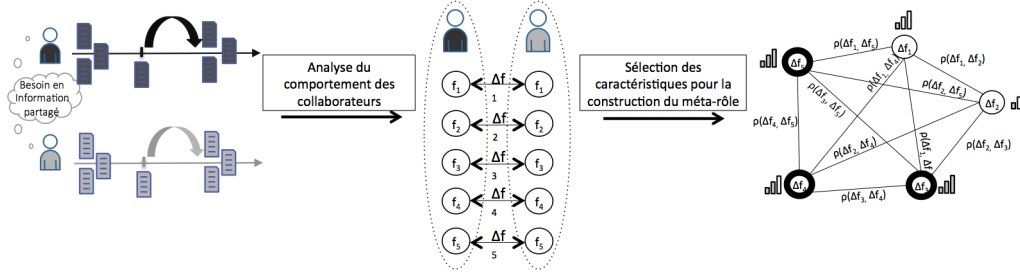


FIGURE 6.2: Méthodologie d'apprentissage des méta-rôles

6.3.1.1 Cadre d'apprentissage des méta-rôles des collaborateurs

Notre méthode d'apprentissage des méta-rôles est inspirée de la méthode de sélection des variables proposée par Geng et al. (2007) et consiste en l'identification du noyau du méta-rôle $\mathcal{K}_{1,2}^{(t_l)}$ qui correspond au plus petit sous-ensemble $F_{sel}^{(t_l)} \in F^p$ de caractéristiques de comportement, où p est non défini, qui répond aux hypothèses suivantes, que nous avons adaptées à notre modèle d'apprentissage du méta-rôle des collaborateurs à l'instant t_l :

- **H1** : l'importance des caractéristiques est dépendante de leur fiabilité à être considérées comme sources d'évidence dans l'assignation des documents aux collaborateurs dans un modèle d'ordonnancement collaboratif. Dans cet objectif, nous définissons une mesure d'importance $\mathcal{A}_{1,2}^{(t_l)}(f_k)$ d'une caractéristique de comportement f_k vis-à-vis des deux collaborateurs u_1 et u_2 à l'instant t_l .
- **H2** : la redondance entre les caractéristiques de comportement doit être évitée dans l'objectif de considérer celles qui sont les plus discriminantes afin de caractériser le méta-rôle au travers de la complémentarité des compétences de recherche des collaborateurs. Ainsi, nous nous appuyons sur la mesure de corrélation $C_{1,2}^{(t_l)}(f_k, f_{k'})$ à partir des caractéristiques de comportement f_k et $f_{k'}$.
- **H3** : la sélection des caractéristiques permettant de faire émerger le méta-rôle doit maximiser l'importance $\mathcal{A}_{1,2}^{(t_l)}(f_k)$ des caractéristiques sélectionnées $f_k \in F_{sel}^{(t_l)}$ dans le modèle d'ordonnancement collaboratif et minimiser la redondance $C_{1,2}^{(t_l)}(f_k, f_{k'})$ entre les paires de caractéristiques sélectionnées.

En tenant compte de l'hypothèse H3, nous formalisons l'algorithme de sélection des caractéristiques pour la construction des méta-rôles selon le modèle

d'optimisation suivant :

$$\begin{aligned}
& \max_{\alpha} \quad \sum_{k=1}^n \mathcal{A}_{1,2}^{(t_l)}(f_k) \cdot \alpha_k \\
& \min_{\alpha} \quad \sum_{k=1}^n \sum_{k'=1}^n C_{1,2}^{(t_l)}(f_k, f_{k'}) \cdot \alpha_k \cdot \alpha_{k'} \\
& \text{sous la contrainte} \quad \alpha_k = \{0, 1\}; \quad k = 1, \dots, n \\
& \text{et} \quad \sum_{k=1}^n \alpha_k = p
\end{aligned} \tag{6.2}$$

avec $\alpha = \{x_1, \dots, x_k, \dots, x_n\}$ est le vecteur de taille n où chaque élément α_k est un indicateur booléen spécifiant si la caractéristique f_k est incluse dans le sous-ensemble de caractéristiques $F_{sel}^{(t_l)}$ à l'instant t_l .

Ce problème d'optimisation à plusieurs objectifs peut être transformé en un problème d'optimisation à unique objectif en combinant linéairement les deux fonctions d'optimisation :

$$\begin{aligned}
& \max_{\alpha} \quad \sum_{k=1}^n \mathcal{A}_{1,2}^{(t_l)}(f_k) \cdot \alpha_k - \gamma (\sum_{k=1}^n \sum_{k'=1}^n C_{1,2}^{(t_l)}(f_k, f_{k'}) \cdot \alpha_k \cdot \alpha_{k'}) \\
& \text{sous la contrainte} \quad \alpha_k = \{0, 1\}; \quad k = 1, \dots, n \text{ et } \sum_{k=1}^n \alpha_k = p
\end{aligned} \tag{6.3}$$

où γ est un paramètre de contrôle.

6.3.1.2 Mesures de qualité des caractéristiques de comportement

Etant donné les hypothèses H1 et H2 de notre problème d'optimisation, nous définissons deux mesures $\mathcal{A}_{1,2}^{(t_l)}(f_k)$ et $C_{1,2}^{(t_l)}(f_k, f_{k'})$, qui traduisent respectivement l'importance et la complémentarité des caractéristiques de comportement dans la session de recherche collaborative.

Importance des caractéristiques de comportement. L'importance d'une caractéristique f_k reflète dans quelle mesure elle constitue un bon indicateur dans l'allocation des documents aux collaborateurs effectuée par un modèle de RIC. Tout comme dans nos précédentes contributions (Soulier et al., 2013, 2014b,c), nous faisons l'hypothèse qu'un modèle collaboratif doit prendre en considération le paradigme de division du travail, et que les documents doivent être assignés au collaborateur qui est le plus à même de les juger. Ainsi, comme proposé par Shah et al. (2010), nous formalisons ce principe au travers d'une classification des documents basée sur les jugements de pertinence effectués par les collaborateurs depuis le début de la session jusqu'à l'instant t_l où l'on souhaite apprendre le méta-rôle. Chaque groupe de documents est associé à un utilisateur et représente les documents alloués à ce dernier. Dans cet objectif, nous proposons de classifier, par le biais d'un algorithme de classification 2-means, l'ensemble $\mathcal{D}^{(t_l)} = \mathcal{D}(u_1)^{(t_l)} \cup \mathcal{D}(u_2)^{(t_l)}$ des documents sélectionnés par les deux collaborateurs u_1 et u_2 jusqu'à l'instant t_l en considérant comme critère de classification la caractéristique f_k . La classe dont le centre de gravité est le plus

élevé est assignée au collaborateur u_j caractérisé par la valeur $S_{u_1}^{(t_l)}(f_k)$ de la caractéristique f_k la plus élevée tandis que l'autre classe est associée à l'autre collaborateur. Dans l'objectif de vérifier la qualité de la caractéristique f_k sur la classification des documents f_k vis-à-vis des deux collaborateurs u_1 et u_2 , nous proposons d'estimer la mesure de fiabilité $\mathcal{A}_{1,2}^{(t_l)}(f_k)$, décrite comme suit et présentée dans l'Illustration 6.1 :

$$\mathcal{A}_{1,2}^{(t_l)}(f_k) = \frac{TP1_{f_k}^{(t_l)} + TP2_{f_k}^{(t_l)}}{TP1_{f_k}^{(t_l)} + FN1^{(t_l)} + TP2_{f_k}^{(t_l)} + FN2^{(t_l)}} \quad (6.4)$$

où $TP1_{f_k}^{(t_l)}$, respectivement $TP2_{f_k}^{(t_l)}$, représente le nombre de documents assignés à l'utilisateur u_1 , respectivement l'utilisateur u_2 , par l'algorithme de classification 2-means basé sur la caractéristique f_k et jugés pertinents par ce même utilisateur à l'instant t_l . $FN1^{(t_l)}$, respectivement $FN2^{(t_l)}$, correspond au nombre de documents non assignés à l'utilisateur u_1 , respectivement à l'utilisateur u_2 , qui a jugé comme pertinents ces mêmes documents à l'instant t_l .

Illustration 6.1 Importance des caractéristiques de comportement.

Nous considérons une session de RIC de 30 minutes impliquant deux utilisateurs u_1 et u_2 afin de résoudre un besoin en information partagé. Cette session est basée sur une collection de 20 documents d_i ; $i = 1 \dots 20$. Considérons qu'à un instant donné $t_l = 300$ (en secondes), l'utilisateur u_1 soumet une requête q . Nous souhaitons identifier le méta-rôle des deux collaborateurs.

Les documents sélectionnés $\mathcal{D}^{(t_l)}(u)$ par chaque collaborateur u sont les suivants :

- $\mathcal{D}^{(t_l)}(u_1) = \{d_1, d_3, d_9\}$, respectivement visités 10, 20 et 15 secondes.
- $\mathcal{D}^{(t_l)}(u_2) = \{d_{12}, d_4, d_6\}$, respectivement visités 2, 8 et 15 secondes.

Nous illustrons le calcul de l'importance des caractéristiques au travers de celle basée sur la durée de visite d'une page web, notée D_p . En réalisant une 2-means sur la base de cette caractéristique, nous obtenons les classes suivantes :

- C_1 : documents $\{d_1, d_{12}, d_4\}$ avec un centre de gravité égal à 6.67.
- C_2 : documents $\{d_3, d_9, d_6\}$ avec un centre de gravité égal à 16.67.

Le collaborateur u_1 a visité les documents en moyenne 15 secondes tandis que l'utilisateur u_2 a visionné les pages 8.33 secondes en moyenne. Par conséquent, la classe C_1 est associée au collaborateur u_2 et, par analogie, la classe C_2 est associée à l'utilisateur u_1 .

La mesure d'importance est donc estimée ainsi :

$$\mathcal{A}_{1,2}^{(300)}(D_p) = \frac{2+2}{2+1+2+1} = 0.67$$

Complémentarité des caractéristiques de comportement. Notre objectif est d’identifier les caractéristiques les plus discriminantes pour caractériser le méta-rôle des collaborateurs à un instant t_l de la session, et plus particulièrement, les caractéristiques qui mettent en évidence des complémentarités dans les compétences de recherche entre collaborateurs. La principale intuition consiste à identifier dans quelles compétences les collaborateurs sont les plus fiables pour résoudre le besoin en information partagé. Pour cela, nous utilisons la mesure de corrélation $C_{1,2}^{(t_l)}(f_k, f_{k'})$ des différences entre collaborateurs $\Delta_{1,2}^{(t_l)}(f_k)$ et $\Delta_{1,2}^{(t_l)}(f_{k'})$ vis-à-vis d’une paire de caractéristiques de comportement f_k and $f_{k'}$. Cette mesure est estimée de façon similaire à celle présentée dans la section 5.3.2 qui présente également un exemple dans l’Illustration 5.4.

6.3.1.3 Optimisation des méta-rôles

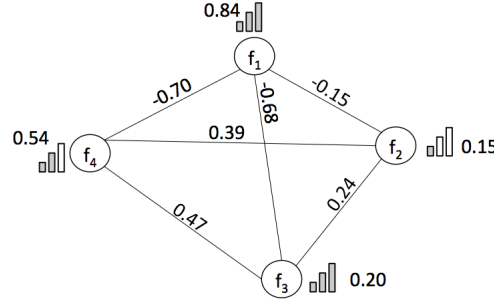
Notre problème d’optimisation défini dans l’Equation 6.3 peut être résolu par une simple exploration de l’ensemble des combinaisons possibles des caractéristiques de comportement de taille μ , où $\mu = 2, \dots, p$. Bien que permettant d’identifier la solution optimale selon une approche exacte, cette méthode peut s’avérer coûteuse en temps avec une complexité pouvant atteindre $O(\sum_{k=1}^p C_n^k)$. Nous proposons dans cette contribution une approche de résolution approximée basée sur la théorie des graphes qui permet d’identifier le meilleur sous-ensemble de caractéristiques.

Dans ce contexte, nous proposons de représenter les caractéristiques de comportement au travers d’un graphe de collaboration $G_{1,2}^{(t_l)}$ qui modélise le comportement des collaborateurs u_1 et u_2 à l’instant t_l . Le graphe $G_{1,2}^{(t_l)} = (A_{1,2}^{(t_l)}, C_{1,2}^{(t_l)})$, illustré dans l’Illustration 6.2, est constitué de noeuds $A_{1,2}^{(t_l)}$ qui représentent chaque caractéristique de comportement $f_k \in F$, pondérée par une mesure d’importance $\mathcal{A}_{1,2}^{(t_l)}(f_k)$ en tant que source d’évidence dans un modèle de RIC, et des relations non orientées $C_{1,2}^{(t_l)} : F \times F$ qui représentent les similarités ou complémentarités des comportements des collaborateurs vis-à-vis des paires de caractéristiques de comportement. Chaque relation $(f_k, f_{k'})$ est pondérée par une mesure de corrélation $C_{1,2}^{(t_l)}(f_k, f_{k'})$ des différences entre caractéristiques de comportement.

Ce dernier peut correspondre à une solution d’optimum local qui est plus applicable compte tenu de sa complexité maximale égale à $O(3^{\frac{n}{3}})$. Le principal objectif de notre méthode consiste en l’extraction du plus petit sous-ensemble de caractéristiques qui maximise les différences entre collaborateurs par rapport à leur comportement et leurs préférences. Cependant, la principale difficulté repose sur la définition du nombre p de caractéristiques à sélectionner. Les notations utilisées sont détaillées dans le Tableau 6.2. Dans ce qui suit, nous présentons l’algorithme, appelé *Coll – Clique* et illustré dans l’Algorithme 5, pour résoudre le problème d’optimisation posé dans la section 6.3.1.3.

Illustration 6.2 Graphe associé au comportement des collaborateurs.

Nous considérons le graphe associé au comportement des collaborateurs u_1 et u_2 associés à la recherche présentée dans l’Illustration 6.1. Nous faisons l’hypothèse que les mesures d’importance et de complémentarité ont été estimées sur 4 caractéristiques $\{f_1, f_2, f_3, f_4\}$. Nous obtenons le réseau suivant :



Par exemple, la valeur 0.84 reliée au noeud f_1 représente la mesure d’importance de la caractéristique f_1 estimée grâce à l’Equation 6.4. La valeur -0.70 de la relation exprimée entre les noeuds f_1 et f_4 correspond à la corrélation entre les différences des collaborateurs vis-à-vis de ces deux caractéristiques, qui traduit le niveau de complémentarité de ces deux caractéristiques de comportement.

Notation	Description
C	Le graphe des caractéristiques représentant la clique en croissance
P	Le graphe candidat évolutif
K	La clique maximale qui satisfait le problème d’optimisation
$Nbhd(C)$	La fonction qui retourne les noeuds caractéristiques, dont le poids est positif, voisins des caractéristiques appartenant à C
$Nodes(K)$	La fonction qui retourne les caractéristiques appartenant à K

TABLE 6.2: Notations utilisées pour l’apprentissage des méta-rôles

L’algorithme *Coll-Clique* est une extension de l’algorithme de clique maximum présenté par Fahle (2002) dans l’objectif d’intégrer le contexte collaboratif dans l’algorithme de sélection des caractéristiques. L’intuition qui nous a guidés est de considérer qu’un graphe pondéré est complet lorsqu’il modélise les complémentarités de comportement au travers des corrélations entre paires de caractéristiques. Par conséquent, plutôt que de se concentrer sur les noeuds du graphe uniquement afin d’identifier le plus grand sous-graphe complet, notre objectif est d’extraire le sous-graphe qui maximise le poids total des noeuds qui le composent, à savoir l’importance $\mathcal{A}_{1,2}^{(t_l)}(f_{k'})$ des caractéristiques de comportement f'_k

Algorithm 4: Algorithme principal

Données: $G_{1,2}^{(t_l)} = (A_{1,2}^{(t_l)}, C_{1,2}^{(t_l)}), \gamma$

Résultat: $F_{sel}^{(t_l)}$

Début

$C = \{\}$

$K = \{\}$

$P = G_{1,2}^{(t_l)}$

$K = Coll - Clique(C, P, \gamma)$

$F_{sel}^{(t_l)} = Nodes(K)$

Retourne $F_{sel}^{(t_l)}$

(cf. Section 6.3.1.3, Hypothèse H1), et de minimiser le poids des relations entre les noeuds du sous-graphe, autrement dit les corrélations $C_{1,2}^{(t_l)}(f_h, f_{k'})$ des paires des différences de caractéristiques f_h et $f_{k'}$ entre collaborateurs u_1 et u_2 (cf. Section 6.3.1.3, Hypothèse H2). L'extraction de sous-graphe répondant à ces deux conditions permet ainsi de répondre au problème d'optimisation (cf. Section 6.3.1.3, Hypothèse H3) formulé dans l'équation 6.3 (cf. Section 6.3.1.3). L'algorithme *Coll-Clique* est récursif et sélectionne une caractéristique f_h à chaque récursion. Dans l'objectif de réduire le poids des caractéristiques similaires, le poids $\mathcal{A}_{1,2}^{(t_l)}(f_{k'})$ des autres caractéristiques $f_{k'}$ est déprécié par la corrélation $C_{1,2}^{(t_l)}(f_h, f_{k'})$ par rapport à la dernière caractéristique sélectionnée f_h .

Plus formellement, l'algorithme *Coll-Clique* repose sur deux graphes de caractéristiques :

1. La clique en croissance C , candidate à être identifiée comme la clique maximale.
2. Le graphe de caractéristiques P , où chacune des caractéristiques est candidate à être ajoutée à la clique en croissance C .

À l'initialisation de l'algorithme, comme illustré dans l'Algorithme 4, C est vide et P est le graphe qui inclut toutes les caractéristiques de comportement, car elles sont a priori toutes candidates à la constitution de la clique en croissance. L'algorithme est lancé sur chaque caractéristique $f_h \in P$ en incrémentant de façon récursive la clique en croissance C , où le poids des caractéristiques f_h est le poids déprécié $\mathcal{A}_{1,2}^{(t_l)}(f_j)$ par rapport à la récursion précédente et les relations avec les autres caractéristiques $f_{k'} \in Nodes(C)$ sont pondérées par les corrélations $C_{1,2}^{(t_l)}(f_h, f_{k'})$. Cette opération est notée $C \oplus f_h$ dans l'Algorithme 5. De plus, l'algorithme crée, au travers de la fonction $Nbhd(C)$, un nouveau graphe de caractéristiques candidates P qui inclut seulement les caractéristiques qui ont un poids positif. Nous définissons $W(K)$ comme la somme des poids des caractéristiques dans la clique maximale K . Nous faisons l'hypothèse que ce poids correspond à l'indicateur que nous souhaitons maximiser dans la mesure où les poids des caractéristiques f_h (mesure d'importance $\mathcal{A}_{1,2}^{(t_l)}(f_h)$) sont dépréciés de

façon récursive par rapport aux corrélations $C_{1,2}^{(t_l)}(f_h, f_{k'})$ vis-à-vis des noeuds adjacents f_k' . Si le poids $W(C) + W(P)$ des caractéristiques incluses dans C et P est plus faible que le poids $W(K)$ dans la clique maximale K identifiée lors de la récursion courante, il n'existe plus de solution pour construire une meilleure clique à partir de C en ajoutant une caractéristique candidate de P avec un poids plus élevé que le poids des caractéristiques de K . A cette condition, la récursion est terminée. Au final, l'ensemble des caractéristiques sélectionnées $F_{sel}^{(t_l)}$ pour la construction des méta-rôles est extraite à partir des noeuds de la clique maximale K .

Un exemple du déroulement de cet algorithme est présenté dans l'illustration 6.3.

Algorithm 5: Coll-Clique

Données: C, P, γ

Résultat: K

Fonction $Coll - Clique(C, P, \gamma)$

Début Pour chaque $f_h \in P$ faire

$$W(C) = \sum_{f_k \in Nodes(C)} \mathcal{A}_{1,2}^{(t_l)}(f_k)$$

$$W(P) = \sum_{f_k \in Nodes(P)} \mathcal{A}_{1,2}^{(t_l)}(f_k)$$

$$W(K) = \sum_{f_k \in Nodes(K)} \mathcal{A}_{1,2}^{(t_l)}(f_k)$$

Si $(W(C) + W(P) \leq W(K))$ **alors**

 /* Retourne la clique maximum */

Retourne K

/* Incrémente la clique en croissance C */

$$C = C \oplus f_h$$

Pour chaque $f_{k'} \in P$ faire

 /* Dépréciation des poids des noeuds */

$$\mathcal{A}_{1,2}^{(t_l)}(f_{k'}) = \mathcal{A}_{1,2}^{(t_l)}(f_{k'}) - C_{1,2}^{(t_l)}(f_h, f_{k'}) * 2\gamma$$

$$P' = Nhd(C)$$

/* Construction de l'ensemble des noeuds candidats */

Si $(P' = \{\})$ **et** $W(C) > W(K)$ **alors**

 /* Enregistre l'optimum local */

$$K = C$$

Si $P' \neq \{\}$ **alors**

 /* Lancement d'une nouvelle récursion */

$Coll - Clique(C, P', \gamma)$

/* Suppression des noeuds pour une nouvelle récursion */

$$C = C \setminus f_h$$

$$P = P \setminus f_h$$

Illustration 6.3 Application de l'algorithme Coll-Clique.

Nous illustrons l'algorithme sur un ensemble de quatre caractéristiques modélisées par le graphe présenté dans l'Illustration 6.2. Nous présentons dans le tableau ci-dessous le début du déroulement de l'algorithme au départ du noeud f_1 . Pour obtenir la clique maximum, il doit être appliqué, de façon analogique, aux autres noeuds du graphe.

Rec.	C	P	K	W(C)	W(P)	W(K)	$C \oplus f_h$	P'
0	-	1 (0.84) 2 (0.3) 3 (0.88) 4 (1.24)	-	-	2.28	-	1 (0.84)	2 (0.3) 3 (0.88) 4 (1.24)
1	1 (0.84)	2 (0.3) 3 (0.88) 4 (1.24)	-	0.84	2.47	-	1 (0.84) 2 (0.3)	3 (0.64) 4 (1.24)
1.1	1 (0.84) 2 (0.3)	3 (0.64) 4 (1.24)	-	1.14	1.88	-	1 (0.84) 2 (0.3) 3 (0.64)	4 (0.07)
1.1.1	1 (0.84) 2 (0.3) 3 (0.64)	4 (0.07)	-	1.78	0.07	-	1 (0.84) 2 (0.3) 3 (0.64) 4 (0.07)	-
1.1.1	1 (0.84) 2 (0.3) 4 (1.24)	3 (0.17)	1-2-3-4	2.38	0.17	1.85	1 (0.84) 2 (0.3) 4 (1.24) 3 (0.17)	-
1.1	1 (0.84) 3 (0.88)	2 (0.06) 4 (0.07)	1-2-3-4	1.72	0.13	1.85	-	-
1.1	1 (0.84) 4 (1.24)	2 (-0.09) 3 (0.41)	1-2-3-4	2.08	0.41	1.85	1 (0.84) 4 (1.24) 3 (0.41)	-

La colonne Rec. exprime le degré de la recursion dans le déroulement de l'algorithme. Les colonnes C, P et K représentent les données d'entrée d'une récursion tandis que les colonnes $C \oplus f_h$ et P' représentent les sorties de la récursion, incluant respectivement la clique en croissance et le nouveau réseau candidat où les poids des caractéristiques ont été dépréciés.

A la fin de cette première série de récursion, la clique maximum locale est constituée des caractéristiques f_1 , f_4 et f_3 compte tenu que le poids des noeuds est supérieur à la clique courante K.

6.3.2 Ordonnancement collaboratif des documents basé sur les méta-rôles

Dans cette section, nous proposons de réinjecter le patron du méta-rôle identifié lors de la section précédente afin d'ordonner collaborativement les documents vis-à-vis des utilisateurs et du besoin en information. L'intuition princi-

pale repose sur l'utilisation des caractéristiques discriminantes, qui modélisent les complémentarités de comportement des collaborateurs, afin (a) d'assigner les documents à l'utilisateur le plus à même de les juger afin de garantir le paradigme de la division du travail et (b) d'ordonner les documents alloués à chaque utilisateur. Dans cet objectif, nous utilisons un algorithme d'apprentissage basé sur un classifieur qui utilise la représentation des documents restreinte aux caractéristiques impliquées dans le noyau $\mathcal{K}_{1,2}^{(t_l)} \in F^p$ du méta-rôle à l'instant t_l . Nous avons choisi d'utiliser le classifieur de régression logistique qui est divisé en deux étapes :

1. L'étape d'apprentissage qui prend en compte l'ensemble $\mathcal{D}^{(t_l)}$ des documents sélectionnés par les deux collaborateurs u_1 et u_2 avant l'instant t_l . Les documents sélectionnés par les deux collaborateurs sont supprimés de cet ensemble, étant donné qu'ils ne sont pas discriminants pour l'allocation des documents collaborative basée sur le paradigme de la division du travail. Chaque document $d_i \in \mathcal{D}^{(t_l)}$ est modélisé par un vecteur de caractéristiques $x_i^{(t_l)} \in \mathcal{R}^p$, estimé à partir des valeurs des caractéristiques pour chacun des documents d_i par rapport aux actions des collaborateurs et de l'instant $t_{l'}$ où le document a été jugé comme pertinent par l'utilisateur, avec $t_{l'} \leq t_l$. Le document d_i reçoit une variable de classification $c_j^{(t_l)} \in \{0; 1\}$ dont la valeur peut être de 0 ou de 1 et exprime la classe associée à chacun des collaborateurs, respectivement u_1 et u_2 , qui ont sélectionné ce document. L'objectif de la fonction d'apprentissage d'ordonnancement des documents identifie le vecteur de poids des prédicteurs $\beta_j^{(t_l)} \in \mathcal{R}^p$ afin d'estimer la probabilité d'assigner un document à un collaborateur $u_j \in \{u_1, u_2\}$. La régression logistique maximise la vraisemblance, détaillée dans l'Equation 6.5 qui repose sur une fonction logit. Cette dernière modélise la probabilité $P_j(x_i^{(t_l)})$ d'un document d_i d'appartenir à une classe utilisateur $c_j^{(t_l)} \in \{0; 1\}$ par rapport à son vecteur de caractéristiques $x_i^{(t_l)}$.

$$\max_{\beta_1^{(t_l)}, \beta_2^{(t_l)}} \sum_{j=1}^2 \sum_{d_i \in \mathcal{D}^{(t_l)}} (c_j \cdot \ln(P_j(x_i^{(t_l)}))) + (1 - c_j) \ln(1 - P_j(x_i^{(t_l)}))$$

$$\text{où } P_j(x_i^{(t_l)}) = \text{logit}(x_i^{(t_l)}) = \frac{\exp(x_i^{(t_l)} \cdot t(\beta_j^{(t_l)}))}{1 + \exp(x_i^{(t_l)} \cdot t(\beta_j^{(t_l)}))} \quad (6.5)$$

où $t(\beta_j^{(t_l)})$ correspond à la transposée du vecteur de coefficients $\beta_j^{(t_l)}$.

2. L'étape de test considère l'ensemble $\mathcal{D}_{ns}^{(t_l)}$ des documents non sélectionnés par les deux collaborateurs u_1 and u_2 avant l'instant t_l . Le vecteur de caractéristiques $x_i^{(t_l)}$ est estimé en fonction des valeurs des caractéristiques pour le document d_i lorsque la valeur est disponible ou sinon selon les valeurs moyennes des caractéristiques des documents à partir des logs de recherche collectés à l'instant t_l , pas nécessairement auprès des deux collaborateurs. En effet, la valeur de la caractéristique "recouvrement de l'annotation" n'est pas disponible pour un document non sélectionné. Elle est donc estimée par la valeur moyenne de cette même caractéristique sur l'ensemble des documents

qui ont été annotés par les utilisateurs du système avant l’instant t_l . A partir de ces représentations des documents non sélectionnés, le modèle appris grâce à la régression logistique estime la probabilité $P_j(x_i^{(t_l)})$ d’assigner un document $d_{i'} \in \mathcal{D}_{ns}^{(t_l)}$ à la classe c_i du collaborateur u_j en tenant compte du vecteur de poids des prédicteurs $\beta_j^{(t_l)}$. Le document $d_{i'}$ est alloué à la classe c_j du collaborateur u_j qui obtient la probabilité $P_j(x_i^{(t_l)})$; $\forall j \in \{0, 1\}$ la plus élevée, qui est également utilisée dans l’ordonnancement des documents au sein même de la classe du collaborateur.

6.4 Expérimentation et résultats

Tout comme notre précédente contribution, nous proposons une évaluation de type *log-study* qui repose sur deux expérimentations utilisateurs. Son objectif consiste en l’analyse de l’impact de notre modèle basé sur les méta-rôles des collaborateurs sur l’efficacité d’une session de RIC. Les hypothèses qui ont guidé nos expérimentations sont les suivantes :

1. Un modèle de RIC doit prendre en compte les compétences et comportements complémentaires des collaborateurs afin de bénéficier de ceux pour lesquels ils sont les plus efficaces.
2. Un modèle de RIC doit être plus efficace qu’un scénario où les utilisateurs recherchent de façon individuelle (effet synergique).
3. Un modèle de RIC doit s’adapter au comportement évolutif des collaborateurs au cours de la session plutôt qu’à des rôles prédéfinis et fixés en début de session.

Dans ce qui suit, nous présentons le protocole expérimental ainsi que les résultats obtenus.

6.4.1 Construction du jeu de données

6.4.1.1 Expérimentations utilisateurs

Nous avons utilisé les logs de recherche de deux expérimentations utilisateurs basées sur la collaboration. Dans l’objectif de mettre en évidence la validité de notre modèle vis-à-vis de jeux de données hétérogènes, nous avons considéré deux expérimentations réalisées de façon indépendante par des acteurs différents, à savoir l’expérimentation *US2* menée au sein de l’Université de Rutgers (USA), comme décrit dans la Section 5.4.1.1, et celle décrite dans la Section 5.1.2.1 que nous avons menées au sein de l’Université Paul Sabatier, que nous avons appelé *SansRole*. Les deux autres scénarios GS et PM ne peuvent être considérés dans la mesure où le comportement des participants est biaisé par la contrainte des rôles. Pour rappel, la tâche de ces deux expérimentations a été d’identifier, durant une session de 30 minutes, le maximum de documents

pertinents en rapport à un besoin exploratoire, à savoir “*global warming*” pour *US2* et “*Tempête tropicale*” pour *SansRole*. Pour cela, les participants ont eu la possibilité de rechercher des informations sur le web en soumettant des requêtes à des moteurs de recherche publics. La collaboration a été supportée par un système de collaboration consistant en un plugin Firefox qui permet aux utilisateurs de sauvegarder des documents, de communiquer et de visualiser l’ensemble de leurs jugements de pertinence. Egalement, ce système a enregistré les actions des participants.

Chaque expérimentation, *SansRole* et *US2*, regroupe respectivement 25 et 10 paires de participants qui ont reçu une compensation pour leur implication dans la tâche de collaboration. Nous soulignons le fait que dans l’expérimentation *SansRole*, les participants sont francophones tandis que les participants de l’expérimentation *US2* sont anglophones. Afin d’éviter un biais dans la tâche de recherche, les participants ont effectué l’expérimentation dans leur langue maternelle, à savoir le français pour *SansRole* et l’anglais pour *US2*.

6.4.1.2 Collection de documents

Compte tenu des deux langues utilisées dans les expérimentations, nous avons construit deux index de documents. Respectivement, pour chaque expérimentation utilisateurs *SansRole* et *US2*, nous avons agrégé (a) les pages visitées par les participants ainsi que (b) les 100 premiers résultats retournés par Google à partir des requêtes soumises par les participants. Chaque page web a été analysée afin d’extraire le contenu textuel présent entre les balises <title> et <p>. Dans l’objectif d’augmenter la taille des documents, nous avons appliqué ce protocole à d’autres expérimentations :

1. Pour l’expérimentation *US2*, nous avons considéré l’expérimentation *US2* présentée dans le chapitre précédent. Nous obtenons ainsi un index de 74 844 pages web.
2. Pour l’expérimentation *SansRole*, nous avons considéré les expérimentations PM et GS, menées en parallèle de cette dernière et décrites dans la section 5.1.2.1. Nous obtenons ainsi un index de 24 226 pages web.

6.4.1.3 Scénarios et modèles de référence

Afin d’analyser notre méthode d’apprentissage des rôles, nous avons considéré deux versions de notre modèle :

1. **MineRank(q)** consiste en l’application, après chaque soumission de requêtes utilisateurs, de notre modèle composé des deux étapes d’apprentissage des méta-rôles et d’ordonnancement collaboratif basé sur les méta-rôles.
2. **MineRank(t)** consiste en l’application de l’étape d’apprentissage des méta-rôles (section 6.3.1) à intervalles de temps t réguliers alors que l’étape d’or-

donnancement collaboratif (section 6.3.1) est effectuée à chaque soumission de requête conformément au patron de méta-rôles détecté au dernier instant t .

Dans l’objectif de tester l’efficacité de notre modèle, nous avons également considéré quatre modèles de référence :

1. **BM25-RIC** fait référence au scénario de recherche individuelle pour lequel le modèle probabiliste BM25 est utilisé à chaque soumission de requête utilisateur. L’objectif de ce scénario est de modéliser une session basée sur une médiation utilisateur où les utilisateurs effectuent la session sur des moteurs de recherche indépendants.
2. **Logit-RIC** fait référence au modèle d’ordonnancement collaboratif qui applique seulement la deuxième étape de notre modèle (Section 6.3.2) sur la base de l’ensemble des caractéristiques de comportement, sans construction des méta-rôles des collaborateurs. Ce scénario modélise une session de type médiation système sans considération des rôles des utilisateurs dans la session.
3. **PM-RIC** fait référence au modèle de RIC basé sur les rôles de Prospecteur –*Prospector*– et de Mineur –*Miner*– où les rôles sont assignés aux utilisateurs sur l’ensemble de la session (Pickens et al., 2008). Ce scénario modélise une session de type médiation système avec rôles prédéfinis.
4. **GS-RIC** fait référence au modèle de RIC basé sur les rôles de Rassembleur –*Gatherer*– et de Surveilleur –*Surveyor*– où l’assignation des rôles est également fixe sur l’ensemble de la session (Shah et al., 2010). Ce scénario modélise également une session de type médiation système avec rôles prédéfinis.

6.4.1.4 Métriques d’évaluation

Les jugements de pertinence (vérité de terrain) pour chacune des expérimentations ont été construits à partir des jugements de pertinence réalisés par les participants, qui englobent leurs actions de sauvegarde, d’annotation ou de snippet. Similairement à la méthode utilisée dans la Section 5.4.1.2, nous avons utilisé une mesure d’accord entre les participants que nous avons renforcé en imposant qu’un document doit être jugé pertinent par deux utilisateurs différents. Au final, nous obtenons un ensemble de 224 et 47 documents pertinents pour les expérimentations *SansRole* et *US2* respectivement.

Dans la continuité des évaluations menées précédemment (Section 5.4.1.6), nous avons considéré les mesures de précision, rappel et F-mesure réalisées au niveau de la session, basées sur les métriques collaboratives définies par (Shah and González-Ibáñez, 2011b), selon le rang $R = 20$. Nous soulignons le fait que, dans l’objectif d’évaluer l’efficacité de notre modèle, les mesures définies par Shah and González-Ibáñez (2011b) sont appliquées aux ordonnancements plutôt qu’aux pages sauvegardées sur l’ensemble de la session.

1. La précision $Prec^S @ R(g)$ au rang R d'un groupe de collaboration g basée sur la session S :

$$Prec^S @ R(g) = \frac{CouvPert(L_S)}{Couverture(L_S)} \quad (6.6)$$

où $CouvPert(L_S)$ correspond au nombre de documents distincts et pertinents dans l'ensemble des listes L_S présentées lors de la session S et $Couverture(L_S)$ correspond au nombre de documents distincts dans l'ensemble des listes L_S présentées lors de la session S .

2. Le rappel $Rappel^S @ R(g)$ au rang R du groupe g basé sur la session S :

$$Rappel^S @ R(g) = \frac{CouvPert(L_S)}{|Pert|} \quad (6.7)$$

où $|Pert|$ représente le nombre de documents pertinents, soit 224 et 47 respectivement pour les expérimentations *SansRole* et *US2*.

3. La F-mesure $F_{syn}^S @ R(g)$ au rang R d'un groupe g basée sur la session S :

$$F^S @ R(g) = \sum_{t=1}^{|T(g)|} \frac{2 * Prec^S @ R(g) * Rappel^S @ R(g)}{Prec^S @ R(g) + Rappel^S @ R(g)} \quad (6.8)$$

6.4.2 Résultats

Dans cette section, nous présentons les différents résultats obtenus lors de notre expérimentation. Compte tenu de la formalisation de notre modèle et des scénarios considérés, nous procédons dans un premier temps à une étape de paramétrage. Ensuite, nous analysons les résultats obtenus en terme d'efficacité selon deux niveaux : session et itération.

6.4.2.1 Paramétrage du modèle de RIC basé sur les méta-rôles

Dans l'objectif de paramétrer notre modèle et de montrer la robustesse de notre modèle vis-à-vis de la tâche collaborative et des sujets, nous effectuons une approche de validation croisée basée sur la maximisation de la F-mesure $F @ 20$ qui consiste en deux étapes, comme illustré dans la Figure 6.3 :

1. L'étape d'apprentissage qui optimise les paramètres du modèle sur l'une des deux collections, par exemple *SansRole*.
2. L'étape d'évaluation qui estime l'efficacité du modèle sur la seconde collection, par exemple *US2*, en considérant les paramètres optimaux obtenus dans l'étape d'apprentissage réalisée sur la première collection, à savoir *SansRole*.

Les deux scénarios considérés pour notre modèle, à savoir $MineRank(q)$ et $MineRank(t)$, nécessitent une étape de paramétrage. Nous présentons ci-dessous les différents paramètres à optimiser et les valeurs optimales obtenues.

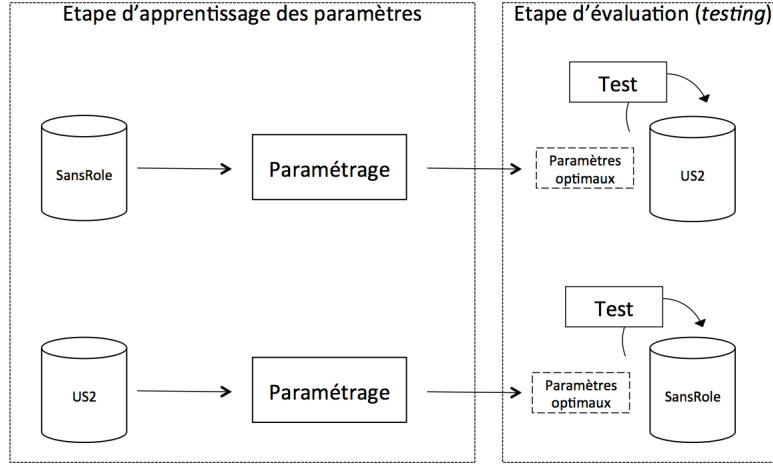


FIGURE 6.3: Protocole d'évaluation basé sur la méthode de validation croisée

Le premier scénario, noté $MineRank(q)$, est lancé à chaque soumission de requête et dépend uniquement du paramètre γ , utilisé dans l'étape d'apprentissage des méta-rôles, comme indiqué dans l'Equation 6.3. Afin de déterminer la valeur optimale, nous avons fait varier sa valeur selon un intervalle de valeurs comprises entre 0 et 1, soit $\gamma \in [0..1]$. Pour chacune de ces valeurs, nous avons ensuite estimé la mesure d'efficacité obtenue pour notre modèle $MineRank(q)$, comme illustré dans la Figure 6.4. Nous pouvons observer que la valeur maximale de la F-mesure est atteinte avec une valeur de 0.05 and 0.13, respectivement pour les expérimentations *SansRole* et *US2*, lorsque le paramètre γ est égal à 0.5 et 0.2, respectivement.

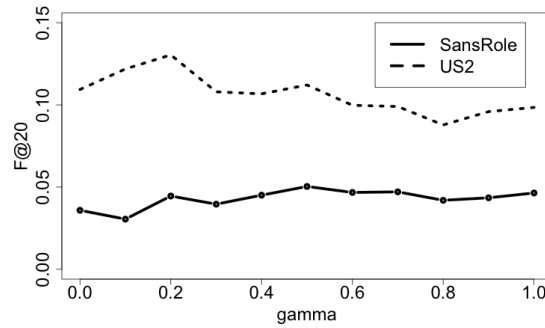


FIGURE 6.4: Paramétrage pour le scénario $MineRank(q)$ selon la valeur γ

Le second scénario de notre modèle, noté $MineRank(t)$, nécessite également de fixer $\gamma \in [0..1]$, comme réalisé précédemment. De plus, compte tenu du cadre d'application où l'apprentissage des rôles est lancé à intervalles réguliers, nous

avons également besoin de déterminer la valeur de ces intervalles t . Pour cela, nous avons considéré une fenêtre temporelle d'apprentissage des rôles entre 1 et 5 minutes dans l'objectif de répondre à notre hypothèse que le comportement des utilisateurs évolue au fur et à mesure de la session. La Figure 6.5(a) et la Figure 6.5(b) illustrent les variations de la F-mesure de notre scénario $MineRank(t)$, respectivement pour les expérimentations *SansRole* et *US2*, compte tenu des deux paramètres $\gamma \in [0..1]$ et $t \in [1..5]$. La F-mesure est optimale ($F@20 = 0.0461$) lorsque $\gamma = 0.5$ et $t = 2$ pour l'expérimentation *SansRole* tandis que pour l'expérimentation *US2* elle atteint 0.112 pour $\gamma = 0.2$ et $t = 2$. La valeur optimale de γ obtenue pour les deux scénarios souligne que la prise en compte des complémentarités des compétences dans l'apprentissage des rôles est plus importante dans l'expérimentation *SansRole* que dans l'expérimentation *US2*.

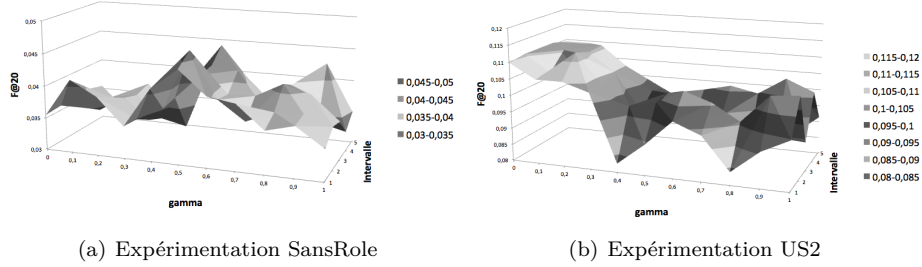


FIGURE 6.5: Paramétrage du scénario $MineRank(t)$

Nous soulignons que les valeurs de F-mesure dans nos deux scénarios, $MineRank(q)$ et $MineRank(t)$, mettent en évidence que l'apprentissage des méta-rôles est plus efficace lorsqu'il est considéré à chaque soumission de requête ($MineRank(q)$) qu'à des intervalles de temps réguliers ($MineRank(t)$). Ce constat confirme notre hypothèse sur le comportement des collaborateurs, qui est consistante avec le fait que les jugements de pertinence sont exprimés après chaque soumission de requête, pas particulièrement reliée avec un intervalle de temps.

6.4.2.2 Analyse de l'efficacité au niveau de la session

Dans cette section, nous comparons l'efficacité de notre approche collaborative basée sur les méta-rôles avec des modèles de l'état de l'art qui considère un scénario individuel, à savoir *BM25-RIC*, un scénario de collaboration sans rôle *Logit-RIC* et deux scénarios de collaboration basés sur des rôles fixes, *PM-RIC* et *GS-RIC*. Le Tableau 6.3 présente les résultats obtenus.

A partir du Tableau 6.3, nous soulignons que, comme observé lors de l'étape de paramétrage, notre modèle est plus efficace lorsqu'il est appliqué à chaque soumission de requête plutôt qu'à des intervalles de temps réguliers. De plus, nous observons que, généralement, notre modèle est plus efficace que l'ensemble des modèles de référence par rapport aux trois mesures d'évaluation. Egalement, nous notons que les améliorations liées à la mesure de rappel sont plus impor-

<i>Apprentissage</i>		<i>Prec^S@20</i>		<i>Rappel^S@20</i>		<i>F^S@20</i>	
→		value	%Tx	value	%Tx	value	%Tx
<i>Evaluation</i>							
<i>US2</i>	BM25-RIC	0,016	+185,64***	0,019	+139,62***	0,0177	+166.71***
	Logit-RIC	0,038	+21,66	0,031	+43.50 *	0,033	+31.75 *
→	GS-RIC	0,015	+204,44 ***	0,008	+429.17 ***	0,009	+345.81 ***
<i>SansRole</i>	PM-RIC	0,019	+136,62 ***	0,006	+719.35 ***	0,008	+432.37 ***
	MineRank(q)	0.046		0.045		0.044	
	MineRank(t)	0.040		0.040		0.040	
<i>SansRole</i>	BM25-RIC	0,075	-5,00	0,063	+335,58	0,069	+63,27
	Logit-CIT	0,071	+0,33	0,266	+3,76 ***	0,111	+1,33 ***
→	GS-RIC	0,058	+23,76	0,039	+595,56 *	0,046	+142,58
<i>US2</i>	PM-RIC	0,092	-22,83	0,078	+254,57	0,084	+32,86
	MineRank(q)	0.071		0.276		0.112	
	MineRank(t)	0.064		0.238		0.112	

TABLE 6.3: Analyse de l'efficacité de notre modèle basé sur l'apprentissage des méta-rôles lors d'une tâche de RIC. %Tx : amélioration de notre modèle *MineRank(q)*. Significativité du test de student * : $0.01 < t \leq 0.05$; ** : $0.001 < t \leq 0.01$; *** : $t \leq 0.001$

tantes que celles liées à la mesure de précision et la F-mesure, ce qui suggère que notre modèle semble fournir des ordonnancements orientés rappel. Plus particulièrement, notre modèle dépasse significativement à la fois les modèles de référence individuels et collaboratifs avec des améliorations comprises entre 43.50% et 719.35% sur l'ensemble des métriques pour l'expérimentation *Sans-Role*. Nous constatons que notre modèle fournit des résultats moins élevés, mais non significatifs que ceux des scénarios *BM25-CIR* et *PM-RIC* pour la mesure de précision. Pour la seconde expérimentation, nous obtenons des améliorations significatives pour le modèle individuel *BM25-RIC* ainsi que pour les modèles collaboratifs *Logit-RIC*, *PM-RIC* and *GS-RIC* pour les trois métriques. Ces résultats soulignent trois conclusions principales :

1. Le méta-rôle des collaborateurs permet d'optimiser la session de recherche lorsqu'il est inféré de façon dynamique et conjointe aux actions des utilisateurs, à savoir la soumission de la requête.
2. Notre modèle permet d'assurer l'effet synergique de la collaboration étant donné que l'efficacité de notre modèle *MineRank(q)* est plus élevée que la somme des résultats individuels, comme illustré dans le scénario *BM25-RIC*.
3. L'apprentissage des méta-rôles semble être plus efficace qu'un scénario de RIC pour lequel les rôles sont explicitement fixés en début de session de recherche, à savoir *GS-RIC* et *PM-RIC*.
4. Les améliorations de notre modèle par rapport au modèle de référence *Logit-RIC* soulignent l'importance de la prise en compte de la complémentarité des compétences des collaborateurs modélisée par l'étape d'apprentissage des méta-rôles et sa réinjection dans le modèle d'ordonnancement collaboratif des documents.

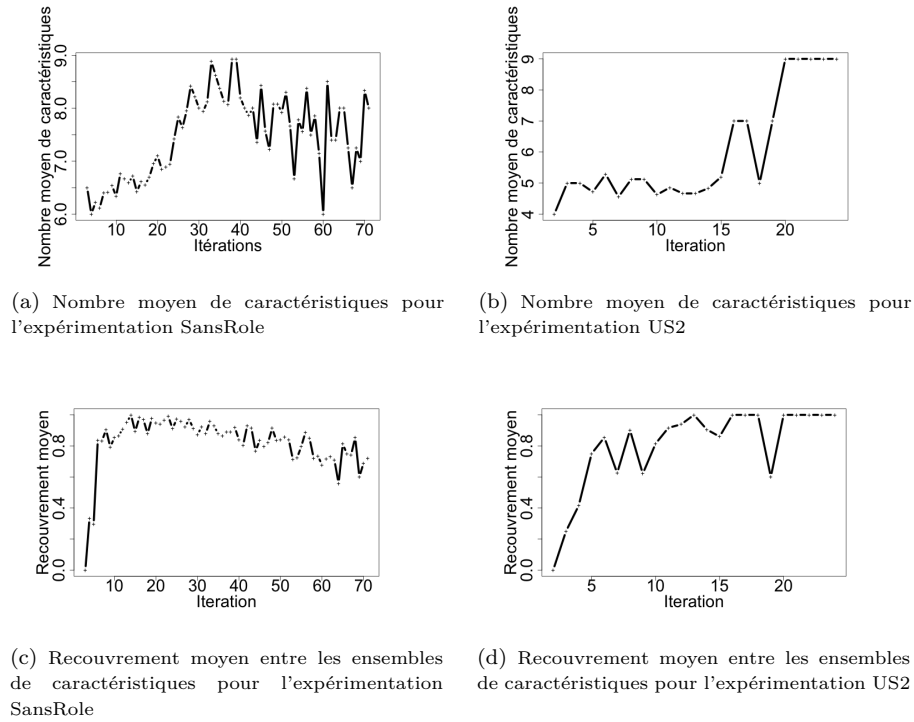


FIGURE 6.6: Analyse temporelle des caractéristiques de comportement impliquées dans l'étape d'ordonnancement collaboratif sur l'ensemble des sessions de recherche

Nous soulignons que, d'après le Tableau 6.3, les taux d'accroissement obtenus pour la collection *SansRole* sont généralement plus élevés que ceux obtenus pour la collection *US2*. Une raison possible qui expliquerait ce constat réside dans le fait que l'apprentissage pour la deuxième collection, *US2* a été réalisé sur une collection plus grande, pouvant conduire à un sur-apprentissage des méta-rôles et du paramètre γ . Dans l'objectif d'analyser cet aspect, nous proposons deux indicateurs que nous estimons à chaque soumission de requête, appelée itération, qui ont été estimés pour chacune des expérimentations comme illustré dans la Figure 6.6 :

1. Le nombre moyen de caractéristiques de comportement retenues pour construire le méta-rôle des collaborateurs.
2. Le recouvrement moyen entre l'ensemble des caractéristiques retenues pour construire les méta-rôles à l'issue de deux requêtes successives.

Chacun des graphiques représente en abscisse les instants de soumission de requêtes, à savoir 71 au total pour l'expérimentation *SansRole* et 24 pour *US2*. Cette différence peut être expliquée par l'objectif supplémentaire d'écrire un rapport qui rassemble les résultats de recherche demandé aux participants de

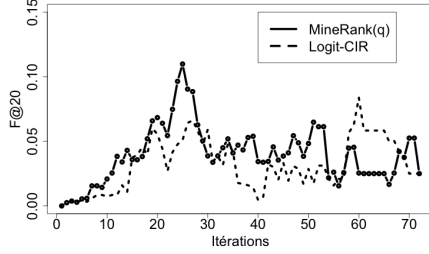
l'expérimentation *US2*, qui a laissé moins de temps aux participants pour naviguer sur le web. Dans l'expérimentation *SansRole*, seulement 3 groupes de collaboration parmi les 25 ont soumis plus de 60 requêtes tandis qu'un seul groupe a soumis plus de 17 requêtes pour la collection *US2*. Ainsi, au delà de cette valeur, les variations identifiées ne sont pas significatives. A partir de la Figure 6.6(a), respectivement la Figure 6.6(b), nous pouvons observer que le nombre de caractéristiques de comportement augmente au fur et à mesure de la session et semble atteindre un maximum entre 8 et 9 caractéristiques, respectivement 7, parmi les 11 considérées dans notre expérimentation. Cette analyse confirme donc notre intuition de sur-apprentissage qui peut être expliquée par le nombre de caractéristiques impliquées dans la construction des méta-rôles.

Dans les Figures 6.6(c) et 6.6(d), l'indicateur de recouvrement est proche de 1 après la 20^{ième} requête soumise pour les deux expérimentations, ce qui souligne que les participants trouvent naturellement le meilleur comportement à adopter compte tenu de leurs compétences, dans l'objectif d'optimiser la session de recherche. Avant cette période, les participants prennent du temps pour s'organiser et identifier les meilleures stratégies à adopter, ce qui fait varier les patrons des méta-rôles entre les requêtes soumises successivement, permettant ainsi, et fur et à mesure, de converger vers un méta-rôle optimal. En conséquence, ces constats donnent de la crédibilité à notre méthodologie d'apprentissage des méta-rôles et souligne la nécessité d'adapter les modèles de RIC au comportement des utilisateurs sur l'ensemble de la session collaborative.

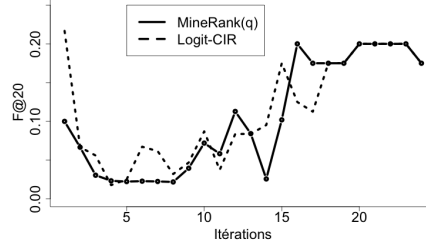
6.4.2.3 Analyse de l'efficacité au niveau de l'itération

Nous complétons notre analyse par une étude de l'efficacité au niveau itératif. Comme indiqué lors de l'étape de paramétrage, nous associons le concept d'itération à chaque instant où un participant soumet une requête. Nous proposons ainsi d'analyser l'efficacité des modèles par rapport aux différentes itérations de la session. Nous estimons l'efficacité d'un groupe g à un instant t_l en considérant les listes de résultats simultanément affichées aux deux utilisateurs à cet instant t_l . Cette métrique correspond à la F-mesure présentée dans la section 5.4.1.6 notée $F^i@R(g)$.

La Figure 6.7 représente l'efficacité au travers de la F-mesure $F@20$ sur l'ensemble de la session à chaque soumission de requête, appelée aussi itération. Pour une meilleure lisibilité, nous avons représenté uniquement le modèle de référence le plus proche, en terme d'efficacité, de notre modèle, soit *Logit-RIC*. Nous observons à partir de la Figure 6.7(a) que la courbe qui représente notre modèle dépasse généralement la courbe du modèle *Logit-RIC* sur l'ensemble de la session. Ce constat est moins net pour l'expérimentation *US2*, du au plus petit nombre de groupes de collaboration. Nous observons deux tendances de courbes différentes selon les expérimentations utilisateurs. En effet, pour l'expérimentation *US1* qui comporte 71 itérations, la courbe d'efficacité décroît en milieu de session à la 30^{ème} itération avant de croître à partir de l'itération 55,



(a) Expérimentation utilisateurs SansRole



(b) Expérimentation utilisateurs US2

FIGURE 6.7: Comparaison de l'efficacité au niveau des itérations sur l'ensemble des sessions

alors que pour l'expérimentation US2, la courbe a une courbe positive à partir de la 2^{ème} itération, et ce, jusqu'à la fin de session. La diminution de la valeur de la F-mesure à l'itération 1 peut être expliquée par la généralité de la première requête soumise par les collaborateurs qui se spécialise par la suite et peut augmenter la difficulté de l'ordonnement associé. Cependant, pour les deux expérimentations, nous observons qu'entre les itérations 15 et 24/30 (suivant les jeux de données), l'efficacité de la session est relativement élevée. Nous pouvons déduire de ce constat que le début de session représente une période latente dans la définition du besoin qui s'affine au fur et à mesure des sessions. Au delà, le nombre de groupes impliqué est plus réduit et donc la consistance des résultats est moins élevée. Du fait de la supériorité globale de la valeur de la mesure d'efficacité sur l'ensemble de la session par rapport au modèle de référence *Logit-RIC*, nous faisons ainsi l'hypothèse qu'une période en début de session est nécessaire (a) à notre modèle afin de réinjecter les jugements de pertinence des participants ainsi qu'aux (b) participants afin d'établir une stratégie de recherche stable qui optimise l'utilisation des compétences de chacun. En effet, ces deux sources d'évidence impactent sur l'apprentissage des méta-rôles, et donc sur l'efficacité de la session et des ordonnancements personnalisés retournés à chaque itération.

Résumé

Dans ce chapitre, nous proposons un modèle non supervisé d'apprentissage dynamique des rôles latents des collaborateurs, appelés également méta-rôles. Tout comme notre précédente contribution (Soulie et al., 2014a), les actions des utilisateurs sont utilisées comme sources d'évidence dans l'analyse de la détection des complémentarités et des similarités du comportement des membres du groupe de collaboration. Cependant, l'atout majeur de cette contribution est qu'elle ne

considère pas de rôles prédéfinis et labellisés, permettant ainsi de définir des ordonnancements qui correspondent aux compétences des collaborateurs, en outre du besoin en information. L'évaluation expérimentale menée sur deux jeux de données issus de scénarios de collaboration réelle prouve l'effet synergique du à notre approche ainsi que sa robustesse du fait de l'utilisation de données multilingues.

Cette contribution comporte cependant certaines limites. Nous avons instancié les besoins des collaborateurs sous la forme d'un modèle d'ordonnement, mais l'identification du méta-rôle, sous-jacente à l'hypothèse de complémentarité des compétences et des stratégies de recherche des collaborateurs, peut conduire à proposer d'autres techniques de RI, telles que la reformulation de requêtes ou l'agrégation des informations, par exemple. De plus, nous avons considéré le cas particulier d'une collaboration synchrone entre deux utilisateurs basée sur un besoin en information exploratoire, à résoudre dans un temps imparti de 30 minutes.

Discussion et conclusion

Dans cette partie, nous nous sommes intéressés particulièrement à la prise en compte des comportements des collaborateurs dans une session de recherche collaborative afin de modéliser leurs rôles ou leurs méta-rôles. Plus particulièrement, nous avons examiné les différences entre collaborateurs dans l'objectif d'identifier les compétences et stratégies de recherche complémentaires pour lesquelles chacun d'eux est le plus efficace. Suite à une analyse préliminaire qui a mis en évidence (a) l'inadéquation des rôles imposés aux collaborateurs par rapport à leurs compétences intrinsèques et, plus particulièrement, relatives au groupe de collaboration et (b) l'évolution des stratégies et compétences de recherche adoptées par le groupe, nous avons proposé de redéfinir une session de recherche collaborative comme une session hybride qui repose à la fois sur :

1. Une médiation utilisateur où les collaborateurs sont libres de s'organiser et de définir les stratégies de recherche ensemble au fur et à mesure de la session.
2. Une médiation système qui guide les collaborateurs dans la lecture des ensembles documentaires en utilisant un modèle d'ordonnancement qui s'adapte aux stratégies de recherche et aux compétences évolutives des collaborateurs.

Dans ce contexte, nous avons proposé d'apprendre les rôles des collaborateurs à partir de leurs actions de façon dynamique au fur et à mesure de la session et de réutiliser ses rôles pour l'ordonnancement des résultats. Similairement à des travaux de l'état de l'art qui analysent le comportement des utilisateurs (Agichtein et al., 2006; White and Dumais, 2009), nous avons modélisé les actions des utilisateurs à partir d'un ensemble de caractéristiques de comportement qui nous a permis d'inférer les compétences et les stratégies de recherche des collaborateurs. Plus particulièrement, notre contribution s'articule autour de deux approches :

1. Un modèle d'apprentissage supervisé (Soulier et al., 2014a) qui a pour objectif d'identifier les rôles des collaborateurs en considérant un ensemble de rôles prédéfinis dans l'état de l'art. L'ordonnancement des documents est réalisé à partir des modèles de RIC (Pickens et al., 2008; Shah et al., 2010) associés aux rôles détectés.
2. Un modèle d'apprentissage non supervisé qui a pour objectif d'identifier les rôles latents des collaborateurs, appelés aussi méta-rôles, afin de les réinjecter dans un modèle d'ordonnancement collaboratif qui alloue les documents aux

collaborateurs en fonction des compétences de recherche pour lesquelles ils sont les plus efficaces.

Les évaluations ont été menées sur des jeux de données réels extraits à partir d'expérimentations utilisateurs réalisées au sein de deux universités. Les résultats ont permis de prouver le bénéfice de notre approche ainsi que de formaliser deux conclusions principales : (a) l'effet synergique de la collaboration est garanti par notre approche dans la mesure où l'efficacité de nos contributions est supérieure à celle d'un scénario de recherche individuelle, (b) l'apprentissage des méta-rôles semble optimiser la collaboration par rapport à un modèle d'ordonnement collaboratif basé seulement sur les caractéristiques de comportement des collaborateurs et (c) la prise en considération de la dynamique des rôles et des actions est efficace par rapport à des scénarios où les rôles sont prédéfinis sur l'ensemble de la session et où le comportement des utilisateurs n'est pas pris en compte.

Bien que soulignant des résultats encourageants de nos deux contributions, nous sommes cependant conscients de leurs limites :

1. Nous avons considéré un scénario de collaboration basé sur une paire d'utilisateurs conformément aux modèles de l'état de l'art basés sur les rôles (Pickens et al., 2008; Shah et al., 2010). Or, de nombreux travaux ont démontré que la plupart des collaborations étaient effectués dans un contexte impliquant 3 à 4 personnes (Morris, 2013). Nous pensons donc qu'il est nécessaire de généraliser ces modèles à un nombre de collaborateurs plus élevé afin de pouvoir être utilisés dans un nombre d'applications plus important.
2. Tout comme dans nos précédents travaux (Soulé et al., 2013, 2014b,c), nous nous sommes basés sur un contexte de collaboration synchrone dont l'objectif est de résoudre une tâche exploratoire. Il nous semble donc important de souligner la limite de portée de nos résultats à d'autres tâches de collaboration.
3. Les expérimentations sont de type *log-study* et ont été réalisées sur des jeux de données réels. Il serait intéressant d'évaluer notre modèle dans un contexte réel où les rôles des participants sont appris en temps réel. Cela permettrait également d'analyser le comportement des utilisateurs face à un tel système et de mesurer la charge cognitive du système.

Quatrième partie

Conclusion générale

Chapitre 7

Conclusion générale

7.1 Synthèse des contributions

Les travaux présentés dans ce manuscrit s'inscrivent dans le contexte de la recherche d'information collaborative qui correspond à un des domaines émergents de la recherche d'information avec de nombreux enjeux multi-disciplinaires tels que la formalisation de modèles d'ordonnancement, l'adaptation des interfaces de collaboration, la prise en compte de l'effort cognitif des utilisateurs et la dualité des interactions entre utilisateurs et avec les sources d'information.

Dans cette thèse, nous nous sommes particulièrement intéressés à la proposition de modèles de recherche d'information collaborative et avons axé notre état de l'art selon cette dimension en donnant un aperçu des différentes approches qui ont été proposées, basées principalement sur les paradigmes de division du travail et de partage des connaissances qui permettent de structurer le processus de collaboration lors d'une tâche de recherche. Une première catégorie de modèles repose principalement sur la prise en compte des jugements de pertinence afin d'ordonner les documents selon un score de pertinence collective ou un score de personnalisation. Cependant, une des principales motivations de la collaboration lors d'une session de recherche est qu'elle doit satisfaire le principe de bénéfice mutuel qui contraint de garantir à la fois la satisfaction du groupe et de l'individu. Pour répondre à cet enjeu, une deuxième approche propose d'intégrer des rôles utilisateurs complémentaires, permettant ainsi d'adapter les ordonnancements aux caractéristiques des rôles tout en favorisant l'effet synergique de la collaboration.

Dans ce contexte, nous avons proposé deux types de contributions, respectivement basées sur le niveau d'expertise des collaborateurs et l'apprentissage dynamique des rôles des collaborateurs dans une session de recherche. Nous les décrivons brièvement ci-dessous :

1. Le niveau d’expertise des collaborateurs est un facteur important dans l’estimation de la pertinence des documents dans la mesure où il permet de personnaliser les résultats afin de permettre à l’utilisateur d’identifier des documents pertinents en fonction de ses compétences et connaissances. Ainsi, nous proposons de considérer l’expertise de domaine selon deux axes : (a) une distinction verticale qui repose sur l’hypothèse qu’un utilisateur est plus expert que son collaborateur et (b) une distinction horizontale où l’hypothèse repose sur le fait que chacun des collaborateurs est expert d’au moins un des sous-domaines du besoin en information. Pour chacune de ces deux distinctions, nous avons proposé un modèle de RIC en deux étapes qui estime dans un premier temps le score des documents vis-à-vis de l’expertise de chaque utilisateur et du besoin en information et, dans un deuxième temps, optimise le score des documents grâce à un algorithme de maximisation de l’espérance afin d’allouer les documents au collaborateur le plus à même de le juger compte tenu de son niveau d’expertise.
2. Grâce à une analyse statistique comparative de sessions de RIC, nous avons mis en évidence le fait que (a) les collaborateurs affinent leurs stratégies de recherche et leur comportement tout au long de la session du fait de leurs interactions avec les sources d’information et les autres membres du groupe et (b) les collaborateurs ont des difficultés à respecter les contraintes d’un rôle imposé et fixé sur l’ensemble de la session. Nous avons ainsi formulé l’hypothèse qu’une session de recherche doit être perçue comme une session hybride qui concilie à la fois une médiation utilisateur et une médiation système afin de laisser les collaborateurs s’organiser librement dans une session de recherche tout en l’optimisant par le biais d’ordonnancements adaptés à leur comportement. Dans cet objectif, nous avons proposé deux contributions basées sur des techniques d’apprentissage des rôles des collaborateurs tout au long de la session de recherche. Plus particulièrement, l’objectif pour chacune de nos contributions est, respectivement, (a) d’inférer de façon supervisée les rôles des collaborateurs sur l’ensemble de la session à partir d’un ensemble de rôles prédéfinis afin d’ordonner les documents selon les modèles de RIC associés et (b) d’inférer les méta-rôles des collaborateurs (rôles latents non prédéfinis) de façon non supervisée afin de les réinjecter dans un modèle d’ordonnement collaboratif.

Ces contributions nous ont permis de poser certains fondements d’un point de vue de l’évaluation de modèles de RIC. En effet, nous avons confirmé d’une part la validité d’un protocole d’évaluation mis en place par Foley and Smeaton (2009) basé sur la simulation de la collaboration en mettant en avant sa flexibilité vis-à-vis des caractéristiques du modèle à évaluer. Par exemple, nous l’avons étendu avec une méthodologie de construction des paires de rôles basée sur le niveau d’expertise des utilisateurs. D’autre part, nous avons appliqué le cadre d’évaluation dans le contexte de la RIC de type *log study* qui permet d’évaluer des modèles sur des jeux de données réelles orientées collaboration. Un autre aspect intéressant que nous avons pu valider tout au long de cette thèse repose sur les métriques d’évaluation que nous avons également adaptées au

contexte collaboratif. Dans leurs précédents travaux, Shah and González-Ibáñez (2011b) proposent des mesures qui sont basées sur les jugements de pertinence des collaborateurs et donc particulièrement orientées pour l’analyse du comportement des utilisateurs et de la production informationnelle du groupe. Dans ce manuscrit, nous proposons d’adapter ces métriques au domaine de la recherche d’information afin de mesurer l’efficacité des modèles de RIC, plutôt que du comportement des utilisateurs. Ainsi, la source d’évidence de la métrique repose sur les listes ordonnées des documents que le modèle fournit aux utilisateurs tout au long de la session, nécessitant de prendre en compte cet aspect interactif dans l’évaluation des modèles de RIC.

7.2 Perspectives

Les différentes évaluations expérimentales menées pour évaluer nos différentes contributions ont montré leur efficacité vis-à-vis des modèles de l’état de l’art ainsi que leur effet synergique par rapport à des scénarios de recherche individuelle. Ce manuscrit ouvre de nombreuses perspectives que nous synthétisons dans cette section.

A court terme, nous souhaitons améliorer nos contributions selon trois aspects :

1. *Modélisation des profils des utilisateurs.* Nous avons proposé d’intégrer les facteurs d’expertise en modélisant l’utilisateur de façon dynamique par l’intermédiaire d’un profil (Chapitre 4). Une des limites de cette modélisation est qu’elle utilise comme sources d’évidence uniquement les actions des collaborateurs au cours de la session, et éventuellement la requête collaborative. Cependant, ces profils ignorent les connaissances et les préférences a priori des utilisateurs et sont uniquement construits sur le court terme. Nous pensons qu’il pourrait être intéressant d’améliorer la modélisation des utilisateurs en prenant en compte ces aspects de personnalisation à la fois sur le court et le long terme.
2. *Estimation de la pertinence collective des documents.* Nos modèles se sont intéressés au calcul de la pertinence des documents en temps réel sur une session de recherche vis-à-vis du besoin en information partagé et des caractéristiques de chacun des collaborateurs selon le principe de division du travail. Afin d’enrichir cette dimension interactive de la pertinence, nous proposons également comme perspective de recherche de construire une fonction d’ordonnancement qui permettrait d’inférer à partir des actions de la session et des caractéristiques des utilisateurs une liste de documents qui traduit, en fin de session de recherche, la pertinence collective des documents. Le défi réside dans la synthèse des actions effectuées par les collaborateurs durant la session de recherche afin d’estimer la pertinence collective des documents vis-à-vis du besoin en information partagé.

A moyen terme, nous proposons de généraliser nos contributions selon deux dimensions, liées aux scénarios de recherche collaborative ainsi qu’au cadre expérimental d’un contexte de RIC :

1. *Généralisation aux différentes tâches de RIC.* Dans la plupart de nos contributions, nous nous sommes intéressés à une paire d’utilisateurs. Une de nos premières perspectives consiste à généraliser ces modèles de RIC pour un groupe de collaboration de taille plus importante afin de répondre à la demande émergente de ces types de groupes (Morris, 2013). De plus, nous nous sommes focalisés sur un type de collaboration synchrone, comme la plupart des travaux de l’état de l’art. Une perspective intéressante serait de proposer des modèles de RIC asynchrone où le facteur temporel ainsi que les interactions entre collaborateurs sont importants dans la prise en compte de la pertinence des documents. Dans ce même esprit, nos contributions ne considèrent en aucun cas les interactions entre utilisateurs qui sont pourtant déterminantes dans un processus de collaboration à la fois au niveau de la formulation du besoin en information partagé que dans l’évaluation de la pertinence des documents. Une analyse intéressante pourrait être d’étudier les relations de leadership dans un groupe de collaboration afin d’identifier si certains individus assument de la pertinence des documents jugés par les autres collaborateurs afin de structurer le processus de collaboration, par exemple.
2. *Cadre formel d’évaluation expérimentale de la RIC.* A ce jour, il n’existe pas de cadre d’évaluation standard permettant d’évaluer les modèles de RIC tel que ceux proposés pour d’autres domaines de la RI par l’intermédiaire des campagnes d’évaluation. Ainsi, une de nos perspectives consiste en la création d’une tâche d’évaluation, du type des tâches TREC. La mise en place d’une telle tâche nécessite la définition/création de plusieurs éléments :
 - (a) Un jeu de données qui recense des sessions de recherche d’information collaborative réelles, comme ceux mis en place lors de nos expérimentations, basées sur plusieurs besoins en information et différentes tâches de recherche (exploratoires, *fact-finding*, ...) afin de répondre aux différents types de collaboration possibles. Plus particulièrement, le jeu de données peut inclure : (i) un ensemble de tâches de collaboration, reliées à des besoins en information partagés, (ii) un ensemble de logs de collaboration réelle qui retracent les actions (requêtes soumises, documents visités, pages sauvegardées...) des collaborateurs impliqués dans la résolution de la tâche, (iii) une collection de documents qui ont permis de réaliser la tâche de collaboration et (iv) un ensemble de documents pertinents qui reflètent la pertinence collective vis-à-vis des tâches de collaboration. Ces derniers pourraient être construits par une méthodologie de pooling où les juges individuels pourraient être remplacés, par exemple, par des groupes de collaboration qui assument de la pertinence graduelle des documents.
 - (b) La définition des mesures d’évaluation orientées collaboration capables de mesurer l’efficacité des modèles de RIC sur la base des ordonnancements de documents. Ces mesures peuvent cependant être affinées

afin de s'approcher au maximum du principe de bénéfice mutuel de la collaboration en proposant des métriques hybrides qui estiment la satisfaction des utilisateurs des ordonnancements vis-à-vis de leurs caractéristiques intrinsèques, sans pour autant dégrader la pertinence collective de la session de recherche. Afin de considérer la difficulté potentielle de fédérer l'ensemble des membres du groupe de collaboration quant à la pertinence des documents, la redéfinition de mesures d'évaluation basées sur la pertinence graduelle, telles que le NDCG, peut être intéressante pour l'évaluation d'une tâche de RIC.

Bibliographie

- Ackerman, M. S. and Palen, L. (1996). The Zephyr Help Instance : Promoting Ongoing Activity in a CSCW System. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '96, pages 268–275. ACM.
- Adcock, J., Pickens, J., Cooper, M. L., Anthony, L., Chen, F., and Qvarfordt, P. (2007). FXPAL Interactive Search Experiments for TRECVID 2007. In *Proceedings of the TRECVID Conference*, TREC '07.
- Agichtein, E., Brill, E., Dumais, S. T., and Ragno, R. (2006). Learning user interaction models for predicting web search result preferences. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Informaion Retrieval*, SIGIR '06, pages 3–10. ACM.
- Allan, J., Aslam, J., Belkin, N., Buckley, C., Callan, J., Croft, B., Dumais, S., Fuhr, N., Harman, D., Harper, D. J., Hiemstra, D., Hofmann, T., Hovy, E., Kraaij, W., Lafferty, J., Lavrenko, V., Lewis, D., Liddy, L., Manmatha, R., McCallum, A., Ponte, J., Prager, J., Radev, D., Resnik, P., Robertson, S., Rosenfeld, R., Roukos, S., Sanderson, M., Schwartz, R., Singhal, A., Smeaton, A., Turtle, H., Voorhees, E., Weischedel, R., Xu, J., and Zhai, C. (2003). Challenges in Information Retrieval and Language Modeling : Report of a Workshop Held at the Center for Intelligent Information Retrieval, University of Massachusetts Amherst, September 2002. *Proceedings of the Forum of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, 37(1) :31–47.
- Allen, B. (1991). Topic knowledge and online catalog search formulation. In *The Library Quarterly*, pages 188–213.
- Amer-Yahia, S., Benedikt, M., and Bohannon, P. (2007). Challenges in Searching Online Communities. *IEEE Data Engineering Bulletin*, 30(2) :23–31.
- Amershi, S. and Morris, M. R. (2008). CoSearch : a system for co-located collaborative web search. In *Proceedings of the Conference on Human Factors in Computing Systems*, CHI '08, pages 1647–1656. ACM.
- Aneiros, M. and Morris, M. R. (2003). Foundation of unconstrained collabora-

- tive web browsing with awareness. In *Proceedings of the International Conference on Web Intelligence*, WI '02, pages 8–25. ACM/IEEE.
- Anick, P. (2003). Using Terminological Feedback for Web Search Refinement : A Log-based Study. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '03, pages 88–95. ACM.
- Assent, I. (2011). Efficient Algorithms for Collaborative Decision Making for Large Scale Settings. In *Proceedings of the International Workshop on Collaborative Information Retrieval*, CIKM '11, pages 15–18. ACM.
- Attfield, S., Adams, A., and Blandford, A. (2006). Patient information needs : pre- and post-consultation. *Health Informatics Journal*, 12 :165–177.
- Attfield, S., Blandford, A., and Makri, S. (2010). Social and interactional practices for disseminating current awareness information in an organisational setting. *Information Processing & Management (IP&M)*, 46(6) :646–655.
- Baeza-Yates, R. and Pino, J. A. (1997). A First Step to Formally Evaluate Collaborative Work. In *Proceedings of the International SIGGROUP Conference on Supporting Group Work : The Integration Challenge*, GROUP '97, pages 56–60. ACM.
- Baeza-Yates, R. A. and Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Addison-Wesley Longman Publishing Co., Inc.
- Balog, K., Azzopardi, L., and de Rijke, M. (2009). A Language Modeling Framework for Expert Finding. *Information Processing & Management (IP&M)*, 45(1) :1–19.
- Beenen, G., Ling, K., Wang, X., Chang, K., Frankowski, D., Resnick, P., and Kraut, R. E. (2004). Using Social Psychology to Motivate Contributions to Online Communities. In *Proceedings of the Conference on Computer Supported Cooperative Work*, CSCW '04, pages 212–221. ACM.
- Belkin, N. J. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science*, (5) :133–143.
- Belkin, N. J. and Croft, W. B. (1992). Information Filtering and Information Retrieval : Two Sides of the Same Coin? *Communications of the ACM (CACM)*, 35(12) :29–38.
- Ben Jabeur, L., Tamine, L., and Boughanem, M. (2010). A social model for literature access : towards a weighted social network of authors. In *Proceedings of the International Conference on Open research Areas in Information Retrieval*, RIAO '10, pages 32–39. ACM.
- Ben Jabeur, L., Tamine, L., and Boughanem, M. (2012). Active Microbloggers : Identifying Influencers, Leaders and Discussers in Microblogging Networks. In *Proceedings of the International Conference on String Processing and Information Retrieval*, SPIRE '12, pages 111–117. Springer.

- Bharat, K. and Mihaila, G. A. (2001). When experts agree : Using non-affiliated experts to rank popular topics. In *Proceedings of the International Conference on World Wide Web*, WWW '01, pages 597–602. ACM.
- Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research (JMLR)*, 3 :993–1022.
- Böhm, T., Klas, C.-P., and Hemmje, M. (2013). Supporting Collaborative Information Seeking and Searching in Distributed Environments. *Proceedings of the "Lernen, Wissen, Adaption" LWA Conference*, pages 16–23.
- Boudighaghen, O., Tamine, L., and Boughanem, M. (2010). A diary study-based evaluation framework for mobile information retrieval. In *Proceedings of the Asia Information Retrieval Societies Conference*, AIRS '10, pages 389–398. Springer.
- Broder, A. (2002). A Taxonomy of Web Search. *Proceedings of the Forum at the Annual International SIGIR Conference on Research and Development in Information Retrieval*, 36(2) :3–10.
- Brousseau-Villeneuve, B., Nie, J., and Kando, N. (2014). Latent word context model for information retrieval. *Information Retrieval*, 17(1) :21–51.
- Buckley, C., Salton, G., and Allan, J. (1994). The Effect of Adding Relevance Information in a Relevance Feedback Environment. In *Proceedings of the Forum at the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '94, pages 292–300. ACM.
- Cao, G., Nie, J., Gao, J., and Robertson, S. (2008). Selecting good expansion terms for pseudo-relevance feedback. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '08, pages 243–250. ACM.
- Capra, R. (2013). Information Seeking and Sharing in Design Teams. In *Proceedings of the ASIS&T Annual Meeting*, ASIS&T '13, pages 239–247. American Society for Information Science.
- Capra, R., Sams, B., and Seligson, P. (2011). Self-Generated Versus Tasks in Collaborative Search. In Science, A. S. f. I., editor, *Proceedings of the International Workshop on Collaborative Information Seeking*, ASIS&T '11. American Society for Information Science.
- Capra, R., Velasco-Martin, J., and Sams, B. (2010). Levels of "Working Together" in Collaborative Information Seeking and Sharing. In *Proceedings of the International Workshop on Collaborative Information Seeking*, CSCW '10. ACM.
- Carbonell, J. and Goldstein, J. (1998). The use of MMR, diversity-based reranking for reordering documents and producing summaries. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '98, pages 335–336. ACM.

- Castells, P., Vargas, S., and Wang, J. (2011). Novelty and Diversity Metrics for Recommender Systems : Choice, Discovery and Relevance. In *Proceedings of the International Workshop on Diversity in Document Retrieval (DDR 2011)*, ECIR '11. Springer.
- Chen, M.-S., Park, J. S., and Yu, P. S. (1998). Efficient Data Mining for Path Traversal Patterns. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 10(2) :209–221.
- Cleverdon, C. (1997). Readings in Information Retrieval. chapter The Cranfi, pages 47–59. Morgan Kaufmann Publishers Inc.
- Clinchant, S. and Gaussier, E. (2010). Information-based models for ad hoc ir. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '10, pages 234–241. ACM.
- Conrad, J. G. (2007). E-Discovery Revisited : A Broader Perspective for IR Researchers. In *Proceedings of the DESI (EDD) Workshop*, ICAIL '07, pages 321–345.
- Cosijn, E. and Ingwersen, P. (2000). Dimensions of relevance. *Information Processing & Management (IP&M)*, 36(4) :533–550.
- Dawes, M. and Sampson, U. (2003). Knowledge management in clinical practice : a systematic review of information seeking behavior in physicians. *International Journal of Medical Informatics*, 71(1) :9–15.
- Demartini, G. (2007). Finding Experts Using Wikipedia. In *Proceedings of the International ExpertFinder Workshop on Finding Experts on the Web with Semantics*, FEWS '07, pages 33–41.
- Demartini, G., Gaugaz, J., and Nejdl, W. (2009). A vector space model for ranking entities and its application to expert search. In *Proceedings of the European Conference on Advances in Information Retrieval*, ECIR '09, pages 189–201. Springer.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1) :1–38.
- Denning, P. J. and Yaholkovsky, P. (2008). Getting to "We". *Communications of the ACM (CACM)*, 51(4) :19–24.
- Deveaud, R., SanJuan, E., and Bellot, P. (2013). Unsupervised Latent Concept Modeling to Identify Query Facets. In *Proceedings of the Open research Areas in Information Retrieval*, OAIR '13, pages 93–100. ACM.
- Diriye, A. and Golovchinsky, G. (2012). Querium : A session-based collaborative search system. In *Proceedings of the European Conference on Advances in Information Retrieval*, ECIR '12, pages 583–584. Springer.

- Dourish, P. and Bellotti, V. (1992). Awareness and coordination in shared workspaces. In *Proceedings of the Conference on Computer Supported Cooperative Work*, CSCW '92, pages 107–114. ACM.
- Dumais, S., Cutrell, E., Cadiz, J. J., Jancke, G., Sarin, R., and Robbins, D. C. (2003). Stuff I’ve Seen : A System for Personal Information Retrieval and Re-use. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '03, pages 72–79. ACM.
- Dumais, S. T. (2009). Evaluating IR in situ. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '09. ACM.
- ECDPC (2011). A literature review on health information-seeking behaviour on the web : a health consumer and health professional perspective. Technical report, European Centre for Disease Prevention and Control.
- Ellis, D. (1989). A Behavioral Approach to Information Retrieval System Design. *Journal of Documentation*, 45(3) :171–212.
- Erickson, T. (2010). A Social Proxy for Collective Search. In *Proceedings of the International Workshop on Collaborative Information Seeking*, CSCW '10. ACM.
- Evans, B. M. and Chi, E. H. (2010). An elaborated model of social search. *Information Processing & Management (IP&M)*, 46(6) :656–678.
- Fahle, T. (2002). Simple and Fast : Improving a Branch-And-Bound Algorithm for Maximum Clique. In *Proceedings of the Annual European Symposium on Algorithms*, ESA' 02, pages 485–498. Springer.
- Fang, H., Tao, T., and Zhai, C. (2004). A formal study of information retrieval heuristics. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '04, pages 49–56. ACM.
- Fidel, R., Pejtersen, A. M., Cleal, B., and Bruce, H. (2004). A Multidimensional Approach to the Study of Human-information Interaction : A Case Study of Collaborative Information Retrieval. *Journal of the Association for Information Science and Technology (JASIST)*, 55(11) :939–953.
- Foley, C. (2008). *Division of labour and sharing of knowledge for synchronous collaborative information retrieval*. PhD thesis, Dublin city University.
- Foley, C. and Smeaton, A. F. (2009). Synchronous Collaborative Information Retrieval : Techniques and Evaluation. In *Proceedings of the European Conference on Advances in Information Retrieval*, ECIR '09, pages 42–53. Springer.
- Foley, C. and Smeaton, A. F. (2010). Division of Labour and Sharing of Knowledge for Synchronous Collaborative Information Retrieval. *Information Processing & Management (IP&M)*, 46(6) :762–772.

- Foster, J. (2006). Collaborative information seeking and retrieval. *Annual Review of Information Science & Technology (ARIST)*, 40(1) :329–356.
- Fox, S. (2008). The engaged E-patient population, Pew internet & American Life Project 2008. Technical report.
- Fox, S., Karnawat, K., Mydland, M., Dumais, S., and White, T. (2005). Evaluating Implicit Measures to Improve Web Search. *ACM Transactions on Information Systems (TOIS)*, 23(2) :147–168.
- Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, 40(1) :35–41.
- Furnas, G. W., Deerwester, S., Dumais, S. T., Landauer, T. K., Harshman, R. A., Streeter, L. A., and Lochbaum, K. E. (1988). Information Retrieval Using a Singular Value Decomposition Model of Latent Semantic Structure. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '88, pages 465–480. ACM.
- Gardner, D. (2005). Ten Lessons in Collaboration. *OJIN : The Online Journal of Issues in Nursing*, 10(1).
- Gauch, S., Chaffee, J., and Pretschner, A. (2003). Ontology-based Personalized Search and Browsing. *Web Intelligence and Agent Systems (WIAS)*, 1(3-4) :219–234.
- Gaver, W. W. (1991). Sound Support for Collaboration. In *Proceedings of the Conference on European Conference on Computer-Supported Cooperative Work*, ECSCW'91, pages 293–308. Kluwer Academic Publishers.
- Geman, S. and Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 6(6) :721–741.
- Geng, X., Liu, T.-Y., Qin, T., and Li, H. (2007). Feature Selection for Ranking. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '07, pages 407–414. ACM.
- Göker, A. and Myrhaug, H. I. (2002). User Context and Personalisation. In *Proceedings of the ECCBR Workshop*, ECCBR '06.
- Golder, S. A. and Donath, J. (2004). Social Roles in Electronic Communities. In *Proceedings of the Association of Internet Researchers Conference*, AoIR '04.
- Golovchinsky, G., Adcock, J., Pickens, J., Qvarfordt, P., and Back, M. (2008). Cerchiamo : a collaborative exploratory search tool. *Proceedings of the Demo in Computer Supported Cooperative Work*.
- Golovchinsky, G., Diriye, A., and Pickens, J. (2011). Designing for Collaboration in Information Seeking. *Proceedings of the ASIS&T Annual Meeting*.

- Golovchinsky, G., Pickens, J., and Back, M. (2009a). A Taxonomy of Collaboration in Online Information Seeking. In *Proceedings of the International Workshop on Collaborative Information Retrieval*, CIR '09.
- Golovchinsky, G., Qvarfordt, P., and Pickens, J. (2009b). Collaborative Information Seeking. *IEEE Computer*, 42(3) :47–51.
- González-Ibáñez, R., Haseki, M., and Shah, C. (2013). Let's search together, but not too close! An analysis of communication and performance in collaborative information seeking. *Information Processing & Management (IP&M)*, 49(5) :1165–1179.
- González-Ibáñez, R., Shah, C., and Cordova-Rubio, N. (2011). Smile! Studying expressivity of happiness as a synergic factor in collaborative information seeking. *Proceedings of the ASIS&T Annual Meeting*, 48(1) :1–10.
- Gray, B. (1989). *Collaborating : finding common ground for multiparty problems*. Jossey Bass Business and Management Series. Jossey-Bass.
- Hansen, P. and Järvelin, K. (2005). Collaborative information retrieval in an information-intensive domain. *Information Processing & Management (IP&M)*, 41(5) :1101–1119.
- Haraty, M., Nobarany, S., and Fisher, B. (2010). ExplorationMap : Supporting Collaborative Exploratory Information Seeking. In *Proceedings of the International Workshop on Collaborative Information Seeking*, CSCW '10. ACM.
- Harman, D. (1992). Relevance Feedback Revisited. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '92, pages 1–10. ACM.
- Hartmann, B., Morris, M. R., Benko, H., and Wilson, A. D. (2009). Augmenting interactive tables with mice & keyboards. In *Proceedings of the Symposium on User Interface Software and Technology*, UIST '09, pages 149–152.
- Harvey, M., Ruthven, I., and Carman, M. (2010). Ranking social bookmarks using topic models. In *Proceedings of the Conference on Information and Knowledge Management*, CIKM '10, pages 1401–1404. ACM.
- Hembroke, H. A., Granka, L. A., Gay, G. K., and Liddy, E. D. (2005). The effects of expertise and feedback on search term selection and subsequent learning. *Journal of the Association for Information Science and Technology (JASIST)*, 56(8) :861–871.
- Hendahewa, C. and Shah, C. (2013). Segmental Analysis and Evaluation of User Focused Search Process. In *Proceedings of the International Conference on Machine Learning and Applications*, ICMLA '13, pages 291–294. IEEE Computer Society.
- Henderson, K., Gallagher, B., Eliassi-Rad, T., Tong, H., Basu, S., Akoglu, L., Koutra, D., Faloutsos, C., and Li, L. (2012). RolX : Structural Role Extraction & Mining in Large Graphs. In *Proceedings of the SIGKDD International*

- Conference on Knowledge Discovery and Data Mining, KDD '12*, pages 1231–1239. ACM.
- Herlocker, J. L., Konstan, J. A., Borchers, A., and Riedl, J. (1999). An Algorithmic Framework for Performing Collaborative Filtering. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '99, pages 230–237. ACM.
- Hertzum, M. (2008). Collaborative information seeking : The combined activity of information seeking and collaborative grounding. *Information Processing & Management (IP&M)*, 44(2) :957–962.
- Heymann, P., Ramage, D., and Garcia-Molina, H. (2008). Social tag prediction. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '08, pages 531–538. ACM.
- Hölscher, C. and Strube, G. (2000). Web search behavior of Internet experts and newbies. *Computer Networks : The International Journal of Computer and Telecommunications Networking*, 33(1-6) :337–346.
- Hopfgartner, F., Vallet, D., Halvey, M., and Jose, J. (2008). Collaborative Search Trails for Video Search. In *Proceedings of the International Workshop on Collaborative Information Retrieval*, JCDL '08. ACM.
- Hupfer, M. E. and Detlor, B. (2006). Gender and Web information seeking : A self-concept orientation model. *Journal of the Association for Information Science and Technology (JASIST)*, 57(8) :1105–1115.
- Hyldegård, J. (2006). Collaborative Information Behaviour : Exploring Kuhlthau's Information Search Process Model in a Group-based Educational Setting. *Information Processing & Management (IP&M)*, 42(1) :276–298.
- Hyldegård, J. (2009). Beyond the Search Process - Exploring Group Members' Information Behavior in Context. *Information Processing & Management (IP&M)*, 45(1) :142–158.
- Imazu, M., Nakayama, S.-i., and Joho, H. (2011). Effect of Explicit Roles on Collaborative Search in Travel Planning Task. In *Proceedings of the Asia Information Retrieval Societies Conference*, AIRS '11, pages 205–214. Springer.
- Ingwersen, P. and Järvelin, K. (2005). Information retrieval in context : IRiX. *Proceedings of the Forum at the Annual International SIGIR Conference on Research and Development in Information Retrieval*, 39(2) :31–39.
- Isenberg, P. and Fisher, D. (2011). Cambiera : collaborative tabletop visual analytics. In *Proceedings of the Conference on Computer Supported Cooperative Work*, CSCW '11, pages 581–582. ACM.
- Jansen, B. J., Booth, D. L., and Spink, A. (2008). Determining the Informational, Navigational, and Transactional Intent of Web Queries. *Information Processing & Management (IP&M)*, 44(3) :1251–1266.

- Järvelin, K. and Kekäläinen, J. (2002). Cumulated Gain-based Evaluation of IR Techniques. *ACM Transactions on Information Systems (TOIS)*, 20(4) :422–446.
- Jelinek, F. and Mercer, R. L. (1980). Interpolated estimation of Markov source parameters from sparse data. In *Proceedings of the Workshop on Pattern Recognition in Practice*, pages 381–397.
- Joachims, T. (2002). Optimizing Search Engines Using Clickthrough Data. In *Proceedings of the SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '02, pages 133–142. ACM.
- Joho, H., Hannah, D., and Jose, J. (2009). Revisiting IR Techniques for Collaborative Search Strategies. In *Proceedings of the European Conference on Advances in Information Retrieval*, ECIR '09, pages 66–77. Springer.
- Joho, H., Hannah, D., and Jose, J. M. (2008). Comparing Collaborative and Independent Search in a Recall-oriented Task. In *Proceedings of the International Symposium on Information Interaction in Context*, IiiX '08, pages 89–96. ACM.
- Kang, I.-H. and Kim, G. (2003). Query Type Classification for Web Document Retrieval. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '03, pages 64–71. ACM.
- Kang, R. and Fu, W.-T. (2010). Exploratory Information Search by Domain Experts and Novices. In *Proceedings of the International Conference on Intelligent User Interfaces*, IUI '10, pages 329–332. ACM.
- Karunakaran, A., Reddy, M. C., and Spence, P. R. (2013). Toward a model of collaborative information behavior in organizations. *Journal of the Association for Information Science and Technology (JASIST)*, 64(12) :2437–2451.
- Kellar, M., Watters, C., and Shepherd, M. (2007). A Field Study Characterizing Web-based Information-seeking Tasks. *Journal of the Association for Information Science and Technology (JASIST)*, 58(7) :999–1018.
- Kelly, D., Dumais, S., and Pedersen, J. O. (2009). Evaluation Challenges and Directions for Information-Seeking Support Systems. *IEEE Computer*, 42(3) :60–66.
- Kelly, R. and Payne, S. J. (2013). Division of labour in collaborative information seeking : Current approaches and future directions. In *Proceedings of the International Workshop on Collaborative Information Seeking*, CSCW '13. ACM.
- Keskustalo, H., Järvelin, K., and Pirkola, A. (2006). The Effects of Relevance Feedback Quality and Quantity in Interactive Relevance Feedback : A Simulation Based on User Modeling. In *Proceedings of the European Conference on Advances in Information Retrieval*, ECIR '06, pages 191–204. Springer.

- Kim, G. (2006). Relationship Between Index Term Specificity and Relevance Judgment. *Information Processing & Management (IP&M)*, 42(5) :1218–1229.
- Kim, H. R. and Chan, P. K. (2003). Learning Implicit User Interest Hierarchy for Context in Personalization. In *Proceedings of the International Conference on Intelligent User Interfaces, IUI '03*, pages 101–108. ACM.
- Kim, K.-S. and Allen, B. (2002). Cognitive and Task Influences on Web Searching Behavior. *Journal of the Association for Information Science and Technology (JASIST)*, 53(2) :109–119.
- Kincaid, J. P., Fishburne, R. P., Rogers, R. L., and Chissom, B. S. (1975). Derivation of new readability formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease formula) for Navy enlisted personnel.
- Kirsch, S., Gnasa, M., and Cremers, A. (2006). Beyond the Web : Retrieval in Social Information Spaces. *ECIR '06*, pages 84–95. Springer.
- Korfatis, N. T., Poulos, M., and Bokos, G. (2006). Evaluating authoritative sources using social networks : an insight from Wikipedia. *Online Information Review*, 30(3) :252–262.
- Korhage, R. (1997). Information storage and retrieval. *John Wiley and Sons Inc.*
- Koutrika, G. and Ioannidis, Y. (2005). A Unified User Profile Framework for Query Disambiguation and Personalization. In *Proceedings of Workshop on New Technologies for Personalized Information Access, PIA '05*, pages 44–53.
- Kraft, R., Maghoul, F., and Chang, C. C. (2005). Y!Q : Contextual Search at the Point of Inspiration. In *Proceedings of the Conference on Information and Knowledge Management, CIKM '05*, pages 816–823. ACM.
- Kuhlthau, C., Spink, A., and Cool, C. (1992). Exploration into stages in the information search process in online information retrieval : communication between users and intermediaries. In *Proceedings of the ASIS&T Annual Meeting, ASIS&T '92*, pages 67–71. American Society for Information Science.
- Kuhlthau, C. C. (1991). Inside the search process : Information seeking from the user's perspective. *Journal of the Association for Information Science and Technology (JASIST)*, 42(5) :361–371.
- Kwak, H., Lee, C., Park, H., and Moon, S. (2010). What is Twitter, a social network or a news media ? In *Proceedings of the International Conference on World Wide Web, WWW '10*, pages 591–600. ACM.
- Lange, L. (1998). Information Seeking by Nurses During Beginning-of-Shift Activities. In Saba, V., Pocklington, D., and Miller, K., editors, *Nursing and Computers, Computers and Medicine*, pages 254–262. Springer.

- Large, A., Beheshti, J., and Rahman, T. (2002). Gender differences in collaborative web searching behavior : an elementary school study. *Information Processing & Management (IP&M)*, 38(3) :427–443.
- Larson, M., Soleymani, M., Serdyukov, P., Rudinac, S., Wartena, C., Murdock, V., Friedland, G., Ordelman, R., and Jones, G. J. F. (2011). Automatic Tagging and Geotagging in Video Collections and Communities. In *Proceedings of the International Conference on Multimedia Retrieval, ICMR '11*, pages 1–8. ACM.
- Lay, K.-L., Malone, T. W., and Yu, K.-C. (1988). Object lens : A “spreadsheet” for cooperative work. *ACM Transactions on Information Systems (TOIS)*, 6(4) :332–353.
- Lee, C.-H., Kim, Y.-H., and Rhee, P.-K. (2001). Web personalization expert with combining collaborative filtering and association rule mining technique. *Expert Systems with Applications (ESWA)*, 21(3) :131–137.
- Leung, K. W.-T., Lee, D. L., Ng, W., and Fung, H. Y. (2012). A Framework for Personalizing Web Search with Concept-based User Profiles. *ACM Transactions on Internet Technology (TOIT)*, 11(4) :1–29.
- Li, Q. and Kim, B. M. (2003). Clustering approach for hybrid recommender system. In *Proceedings of the International Conference on Web Intelligence, WI '03*, pages 33–38. IEEE Computer Society.
- Li, X., Guo, L., and Zhao, Y. E. (2008). Tag-based Social Interest Discovery. In *Proceedings of the International Conference on World Wide Web, WWW '08*, pages 675–684. ACM.
- Liang, F., Qiang, R., and Yang, J. (2012). Exploiting real-time information retrieval in the microblogosphere. In *Proceedings of the Joint conference on Digital Libraries, JCDL '12*, pages 267–276. ACM.
- Lieberman, H. (1997). Autonomous Interface Agents. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '97*, pages 67–74. ACM.
- Liechti, O. and Sumi, Y. (2002). Editorial : Awareness and the WWW. *International Journal of Human Computer Studies*, 56(1) :1–5.
- Liu, D., Hua, X.-S., Wang, M., and Zhang, H.-J. (2010). Image Retagging. In *Proceedings of the International Conference on Multimedia, MM '10*, pages 491–500. ACM.
- Liu, F., Yu, C., and Meng, W. (2004). Personalized Web Search For Improving Retrieval Effectiveness. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 16(1) :28–40.
- London, S. (1995). Collaboration and community.

- Low, Y., Agarwal, D., and Smola, A. J. (2011). Multiple Domain User Personalization. In *Proceedings of the SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, pages 123–131. ACM.
- Ma, H., Lyu, M. R., and King, I. (2009). Learning to Recommend with Trust and Distrust Relationships. In *Proceedings of the Conference on Recommender Systems*, RecSys '09, pages 189–196. ACM.
- MacKay, D. J. C. and Peto, L. C. B. (1994). A Hierarchical Dirichlet Language Model. *Natural Language Engineering*, 1(3) :1–19.
- Maisonnasse, L., Gaussier, É., and Chevallet, J. (2007). Revisiting the dependence language model for information retrieval. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '07, pages 695–696. ACM.
- Malik, S., Tombros, A., and Larsen, B. (2007). The Interactive Track at INEX 2006. In Fuhr, N., Lalmas, M., and Trotman, A., editors, *Comparative Evaluation of XML Information Retrieval Systems*, volume 4518 of *Lecture Notes in Computer Science*, pages 387–399. Springer Berlin Heidelberg.
- Manmatha, R., Rath, T., and Feng, F. (2001). Modeling Score Distributions for Combining the Outputs of Search Engines. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '01, pages 267–275. ACM.
- Marchionini, G. (1995). *Information Seeking in Electronic Environments*. Cambridge University Press.
- Marsland, S. (2009). *Machine Learning : An Algorithmic Perspective*. Chapman & Hall/CRC.
- Martin, I. and Jose, J. M. (2004). Fetch : A Personalised Information Retrieval Tool. In *Proceedings of the International Conference on Open research Areas in Information Retrieval*, RIAO '04, pages 405–419. ACM.
- Mascaro, C. M. and S.P., G. (2010). Collaborative Information Seeking in an On-line Political Group Environment. In *Proceedings of the International Workshop on Collaborative Information Seeking*, CSCW '10. ACM.
- McCallum, A., Wang, X., and Corrada-Emmanuel, A. (2007). Topic and Role Discovery in Social Networks with Experiments on Enron and Academic Email. *Journal on Artificial Intelligence Research (JAIR)*, 30(1) :249–272.
- McMullan, M. (2006). Patients using the Internet to obtain health information : How this affects the patient-health professional relationship. *Patient Education and Counseling*, 63(1–2) :24–28.
- Micarelli, A. and Sciarrone, F. (2004). Anatomy and Empirical Evaluation of an Adaptive Web-Based Information Filtering System. *User Modeling and User-Adapted Interaction (UMUAI)*, 14(2-3) :159–200.

- Mimno, D. and McCallum, A. (2007). Expertise Modeling for Matching Papers with Reviewers. In *Proceedings of the SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '07, pages 500–509. ACM.
- Mooney, R. J. and Roy, L. (2000). Content-based Book Recommending Using Learning for Text Categorization. In *Proceedings of the Conference on Digital Libraries*, DL '00, pages 195–204. ACM.
- Moraveji, N., Morris, M., Morris, D., Czerwinski, M., and Henry Riche, N. (2011). ClassSearch : facilitating the development of web search skills through social learning. In *Proceedings of the Conference on Human Factors in Computing Systems*, CHI '11, pages 1797–1806. ACM.
- Morris, M. R. (2013). Collaborative Search Revisited. In *Proceedings of the Conference on Computer Supported Cooperative Work*, CSCW '13, pages 1181–1192. ACM.
- Morris, M. R. and Horvitz, E. (2007). SearchTogether : an interface for collaborative web search. In *Proceedings of the Symposium on User Interface Software and Technology*, UIST '07, pages 3–12.
- Morris, M. R., Lombardo, J., and Wigdor, D. (2010). WeSearch : supporting collaborative search and sensemaking on a tabletop display. In *Proceedings of the Conference on Computer Supported Cooperative Work*, CSCW '10, pages 401–410. ACM.
- Morris, M. R. and Morris, D. (2011). Understanding the Potential for Collaborative Search Technologies in Clinical Settings. In *Proceedings of the International Workshop on Collaborative Information Retrieval*, CIKM '11. ACM.
- Morris, M. R., Paepcke, A., and Winograd, T. (2006). TeamSearch : Comparing Techniques for Co-Present Collaborative Search of Digital Media. In *Proceedings of the International Workshop on Horizontal Interactive Human-Computer Systems*, Tabletop '06, pages 97–104. IEEE Computer Society.
- Morris, M. R. and Teevan, J. (2009). *Collaborative Web Search : Who, What, Where, When, and Why*. Synthesis Lectures on Information Concepts, Retrieval, and Services. Morgan & Claypool Publishers.
- Morris, M. R., Teevan, J., and Bush, S. (2008). Collaborative Web Search with Personalization : Groupization, Smart Splitting, and Group Hit-highlighting. In *Proceedings of the Conference on Computer Supported Cooperative Work*, CSCW '08, pages 481–484. ACM.
- Na, S.-H., Kang, I.-S., and Lee, J.-H. (2008). Improving term frequency normalization for multi-topical documents and application to language modeling approaches. In *Proceedings of the European Conference on Advances in Information Retrieval*, ECIR'08, pages 382–393. Springer.

- Newman, M. E. J. and Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, 69(2) :26113.
- Nowicki, K. and Snijders, T. A. B. (2001). Estimation and prediction for stochastic blockstructure. *Journal of the Association for Information Science and Technology (JASIST)*, 96(455) :1077–1087.
- Oard, D. W. (2001). Modeling Information Content Using Observable Behavior. In *Proceedings of the ASIS&T Annual Meeting*, ASIS&T '01, pages 38–45. American Society for Information Science.
- Oeldorf-Hirsch, A., Hecht, B., Morris, M. R., Teevan, J., and Gergle, D. (2014). To Search or to Ask : The Routing of Information Needs Between Traditional Search Engines and Social Networks. In *Proceedings of the Conference on Computer Supported Cooperative Work*, CSCW '14, pages 16–27. ACM.
- Ogawa, Y., Morita, T., and Kobayashi, K. (1991). A Fuzzy Document Retrieval System Using the Keyword Connection Matrix and a Learning Method. *Fuzzy Sets and Systems*, 39(2) :163–179.
- Osman, D., Yearwood, J., and Vamplew, P. (2010). Automated opinion detection : Implications of the level of agreement between human raters. *Information Processing & Management (IP&M)*, 46(3) :331–342.
- Over, P. (2001). The TREC interactive track : an annotated bibliography. *Information Processing & Management (IP&M)*, 37(3) :369–381.
- Pal, A. and Counts, S. (2011). Identifying topical authorities in microblogs. In *Proceedings of the Conference on Web Search and Data Mining*, WSDM '11, pages 45–54. ACM.
- Park, T. K. (1994). Toward a Theory of User-based Relevance : A Call for a New Paradigm of Inquiry. *Journal of the Association for Information Science and Technology (JASIST)*, 45(3) :135–141.
- Patel, V. L. and Arocha J.F., K. D. R. (1999). *Tacit Knowledge in Professional Practice*, chapter Expertise, pages 75–99. Jossey-Bass Publishers.
- Paul, S. A. and Morris, M. R. (2009). CoSense : Enhancing Sensemaking for Collaborative Web Search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 1771–1780. ACM.
- Paul, S. A. and Reddy, M. C. (2010). A Framework for Sensemaking in Collaborative Information Seeking. In *Proceedings of the International Workshop on Collaborative Information Seeking*, CSCW '10. ACM.
- Pessiot, J., Truong, T., Usunier, N., Amini, M., and Gallinari, P. (2007). Filtrage collaboratif avec un algorithme d'ordonnement. In *Actes de la COnférence en Recherche d'Informations et Applications*, pages 165–180. Université de Saint-Étienne.

- Petkova, D. and Croft, W. B. (2006). Hierarchical Language Models for Expert Finding in Enterprise Corpora. In *Proceedings of the International Conference on Tools with Artificial Intelligence*, ICTAI '06, pages 599–608. IEEE Computer Society.
- Pickens, J., Golovchinsky, G., Shah, C., Qvarfordt, P., and Back, M. (2008). Algorithmic Mediation for Collaborative Exploratory Search. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '08, pages 315–322. ACM.
- Pirkola, A. and Järvelin, K. (2001). Employing the resolution power of search keys. *Journal of the Association for Information Science and Technology (JASIST)*, 52(7) :575–583.
- Podichetty, V. K., Booher, J., Whitfield, M., and Biscup, R. S. (2006). Assessment of internet use and effects among healthcare professionals : a cross sectional survey. *Postgraduate Medical Journal*, 82(966) :274–279.
- Poltrack, S., Grudin, J., Dumais, S., Fidel, R., Bruce, H., and Pejtersen, A. M. (2003). Information Seeking and Sharing in Design Teams. In *Proceedings of the International SIGGROUP Conference on Supporting Group Work*, GROUP '03, pages 239–247. ACM.
- Ponte, J. M. and Croft, W. B. (1998). A language modeling approach to information retrieval. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '98, pages 275–281. ACM.
- Porter, M. F. (1997). Readings in Information Retrieval. chapter An Algorit, pages 313–316. Morgan Kaufmann Publishers Inc.
- Privault, C., O'Neill, J., Ciriza, V., and Renders, J.-M. (2010). A new tangible user interface for machine learning document review. *Artificial Intelligence and Law*, 18(4) :459–479.
- Radinsky, K., Svore, K. M., Dumais, S. T., Shokouhi, M., Teevan, J., Bocharov, A., and Horvitz, E. (2013). Behavioral Dynamics on the Web : Learning, Modeling, and Prediction. *ACM Transactions on Information Systems (TOIS)*, 31(3) :1–37.
- Rafrafi, A., Guigue, V., and Gallinari, P. (2012). Coping with the document frequency bias in sentiment classification. In *Proceedings of the International Conference on Weblogs and Social Media*, pages 314–321. The AAAI Press.
- Reddy, M. C., Pratt, W., Dourish, P., Pratt, A., Michael, M., and Shabot, M. M. (2002). Asking Questions : Information Needs in a Surgical Intensive Care Unit. In *Proceedings of the American Medical Informatics Association Fall Symposium*, AMIA '02, pages 651–655. AMIA.
- Reddy, M. C. and Spence, P. R. (2002). Finding answers : information needs of a multidisciplinary patient care team in an emergency department. In

- Proceedings of the American Medical Informatics Association Fall Symposium*, AMIA '02, pages 649–653. AMIA.
- Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., and Riedl, J. (1994). GroupLens : An Open Architecture for Collaborative Filtering of Netnews. In *Proceedings of the Conference on Computer Supported Cooperative Work*, CSCW '94, pages 175–186. ACM.
- Ringel, M., Cutrell, E., Dumais, S. T., and Horvitz, E. (2003). Milestones in Time : The Value of Landmarks in Retrieving Information from Personal Stores. In *Proceedings of the International Conference on Human-Computer Interaction*, INTERACT '03, pages 184–191. IOS Press.
- Robertson, S. E. and Walker, S. (1994). Some simple effective approximations to the 2-Poisson model for probabilistic weighted retrieval. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '94, pages 232–241. ACM.
- Robertson, S. E., Walker, S., Jones, S., Hancock-Beaulieu, M., and Gatford, M. (1995). Okapi at TREC-3. In *Proceedings of the Text retrieval conference-3 (TREC-3)*, TREC '95, pages 109–126.
- Rocchio, J. J., editor (1971). *Relevance Feedback in Information Retrieval*. Prentice Hall.
- Rodriguez Perez, J., Whiting, S., and Jose, J. M. (2011). CoFox : A visual collaborative browser. In *Proceedings of the International Workshop on Collaborative Information Retrieval*, CIKM '11. ACM.
- Rose, D. E. and Levinson, D. (2004). Understanding User Goals in Web Search. In *Proceedings of the International Conference on World Wide Web*, WWW '04, pages 13–19. ACM.
- Rudd, J. and Rudd, M. J. (1986). Coping with information load : strategies and implications for librarians. *College & Research Libraries*, 47(4) :315–322.
- Sakagami, H. and Kamba, T. (1997). Learning personal preferences on online newspaper articles from user behaviors. *Computer Networks and ISDN Systems*, 29(8–13) :1447–1455.
- Salton, G. (1971). A comparison between manual and automatic indexing method. *Journal of American Documentation*, 2(1) :61–71.
- Salton, G. and Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management (IP&M)*, 24(5) :513–523.
- Salton, G. and McGill, M. J. (1983). *Introduction to modern information retrieval*. McGraw-Hill computer science series. McGraw-Hill.
- Salton, G., Wong, A., and Yang, C. S. (1975). A Vector Space Model for Automatic Indexing. *Communications of the ACM (CACM)*, 18(11) :613–620.

- Samuels, M. A. (2011). The importance of collaboration among physicians. *Journal of Archives of Internal Medicine*, 171(14) :1301.
- Sarcevic, A. (2009). A Study of Collaboration Information Behavior in Trauma Resuscitation.
- Sebastiani, F. (2002). Machine Learning in Automated Text Categorization. *ACM Computing Surveys*, 34(1) :1–47.
- Shah, C. (2012). Collaborative Information Seeking - The Art and Science of Making the Whole Greater than the Sum of All. pages I–XXI, 1–185.
- Shah, C. (2013). Collaborative information seeking (cis) : Challenges and opportunities. In *Proceedings of the International Workshop on Collaborative Information Seeking, CSCW '13*. ACM.
- Shah, C. and González-Ibáñez, R. (2011a). Coagmento - A System for Supporting Collaborative Information Seeking. In *Demo in Proceedings of Association for Information Science and Technology Annual Meeting, ASIST '12*, pages 9–12.
- Shah, C. and González-Ibáñez, R. (2011b). Evaluating the Synergic Effect of Collaboration in Information Seeking. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval, SIGIR '11*, pages 913–922. ACM.
- Shah, C., Pickens, J., and Golovchinsky, G. (2010). Role-based results redistribution for collaborative information retrieval. *Information Processing & Management (IP&M)*, 46(6) :773–781.
- Shen, X., Tan, B., and Zhai, C. (2005). Context-sensitive Information Retrieval Using Implicit Feedback. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval, SIGIR '05*, pages 43–50. ACM.
- Sieg, A., Mobasher, B., and Burke, R. (2007). Web Search Personalization with Ontological User Profiles. In *Proceedings of the Conference on Information and Knowledge Management, CIKM '07*, pages 525–534. ACM.
- Sieg, A., Mobasher, B., Lytinen, S., and Burke, R. (2004). Using concept hierarchies to enhance user queries in web-based information retrieval. In *Proceedings of the International Conference on Artificial Intelligence and Applications, IAA '04*.
- Singer, G., Danilov, D., and Norbistrath, U. (2012). Complex search : Aggregation, discovery, and synthesis. volume 61, pages 89–106. Estonian Academy publishers.
- Smeaton, A. F., Foley, C., Gurrin, C., Lee, H., and McGivney, S. (2006). Collaborative Searching for Video Using the Fischlar System and a DiamondTouch

- Table. In *Proceedings of the International Workshop on Horizontal Interactive Human-Computer Systems*, Tabletop '06, pages 151–159. IEEE Computer Society.
- Smyth, B., Balfe, E., Boydell, O., Bradley, K., Briggs, P., Coyle, M., and Freyne, J. (2005). A live-user evaluation of collaborative web search. In *Proceedings of the International Joint Conference on Artificial Intelligence*, IJCAI '05, pages 1419–1424.
- Soboroff, I. and Harman, D. (2005). Novelty Detection : The TREC Experience. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, HLT '05, pages 105–112. Association for Computational Linguistics.
- Song, F. and Croft, W. B. (1999). A General Language Model for Information Retrieval. In *Proceedings of the Conference on Information and Knowledge Management*, CIKM '99, pages 316–321. ACM.
- Sonnenwald, D. H., Maglaughlin, K. L., and Whitton, M. C. (2004). Designing to support situation awareness across distances : an example from a scientific collaboratory. *Information Processing & Management (IP&M)*, 40(6) :989–1011.
- Soulier, L., Shah, C., and Tamine, L. (2014a). User-driven System-mediated Collaborative Information Retrieval. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '14, pages 485–494. ACM.
- Soulier, L., Tamine, L., and Bahsoun, W. (2013). A Collaborative Document Ranking Model for a Multi-faceted Search. In *Proceedings of the Asia Information Retrieval Societies Conference*, AIRS '13, pages 109–120. Springer.
- Soulier, L., Tamine, L., and Bahsoun, W. (2014b). On domain expertise-based roles in collaborative information retrieval. *Information Processing & Management (IP&M)*, 50(5) :752–774.
- Soulier, L., Tamine, L., and Bahsoun, W. (2014c). Un modèle de recherche d'information collaborative basé sur l'expertise des utilisateurs. In *Actes de la Conférence en Recherche d'Information et Applications*, CORIA '14, pages 35–50. LORIA.
- Spence, P. R., Reddy, M. C., and Hall, R. (2005). A Survey of Collaborative Information Seeking Practices of Academic Researchers. In *Proceedings of the International SIGGROUP Conference on Supporting Group Work*, GROUP '05, pages 85–88. ACM.
- Strijbos, J.-W., Martens, R. O. B. L., Jochems, W. M. G., and Broers, N. J. (2004). The Effect of Functional Roles on Group Efficiency. Using Multilevel Modeling and Content Analysis to Investigate Computer-Supported Collaboration in Small Groups. *Journal of Information Science (JIS)*, 35(2) :195–229.

- Surowiecki, J. (2004). *The Wisdom of Crowds : Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. Doubleday.
- Symeonidis, P., Nanopoulos, A., and Manolopoulos, Y. (2008). Tag Recommendations Based on Tensor Dimensionality Reduction. In *Proceedings of the Conference on Recommender Systems*, RecSys '08, pages 43–50. ACM.
- Talja, S. (2002). Information sharing in academic communities : Types and levels of collaboration in information seeking and use. *New Review of Information Behaviour Research*, 3 :143–160.
- Tamine-Lechani, L. and Bahsoun, W. (2006). Définition d'un profil multidimensionnel de l'utilisateur : Vers une technique basée sur l'interaction entre dimensions. In *COnférence en Recherche d'Informations et Applications*, pages 225–236. Lyon.
- Tamine-Lechani, L., Boughanem, M., and Daoud, M. (2010). Evaluation of contextual information retrieval effectiveness : overview of issues and research. *Knowledge and Information Systems (KAIS)*, 24(1) :1–34.
- Tao, Y. and Tombros, A. (2013). An Exploratory Study of Sensemaking in Collaborative Information Seeking. In *Proceedings of the European Conference on Advances in Information Retrieval*, ECIR '13, pages 26–37. Springer.
- Taylor, R. S. (1968). Question-negotiation and information seeking in libraries. *College & Research Libraries*, 29(3) :178–194.
- Teevan, J., Dumais, S. T., and Horvitz, E. (2005). Personalizing Search via Automated Analysis of Interests and Activities. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '05, pages 449–456. ACM.
- Teevan, J., Morris, M. R., and Bush, S. (2009). Discovering and using groups to improve personalized search. In *Proceedings of the International Conference on Web Search and Data Mining*, WSDM '09, pages 15–24. ACM.
- Thomson, A. M. and Perry, J. L. (2006). Collaboration Processes : Inside the Black Box. *Public Administration Review*, 66 :20–32.
- Tungare, M., Hanrahan, B., Quintana-Castillo, R., Stewart, M., and Perez-Quinones, M. A. (2010). Collaborative human Computation as a Mean of Information Management. In *Proceedings of the International Workshop on Collaborative Information Seeking*, CSCW '10. ACM.
- Twidale, M. B., Nichols, D. M., and Paice, C. D. (1997). Browsing is a Collaborative Process. *Information Processing & Management (IP&M)*, 33(6) :761–783.
- Twidale, M. B., Nichols, D. M., Smith, G., and Trevor, J. (1995). Supporting Collaborative Learning During Information Searching. In *Proceedings of the International Conference on Computer Support for Collaborative Learning*, CSCL '95, pages 367–370. L. Erlbaum Associates Inc.

- Vivian, R. and Dinet, J. (2008). RCI WEB : un système collaboratif de recherche d'information centré utilisateur. *Revue des Interactions Humaines Médiatisées*, 9(2) :85–110.
- Wald, H. S., Dube, C. E., and Anthony, D. C. (2007). Untangling the Web—The impact of Internet use on health care and the physician–patient relationship. *Patient Education and Counseling*, 68(3) :218–224.
- Wang, J. and Soergel, D. (2010). A User Study of Relevance Judgments for e-Discovery. In *Proceedings of the ASIS&T Annual Meeting*, ASIS&T '10, pages 74 :1–74 :10. American Society for Information Science.
- Welser, H. T., Gleave, E., Fisher, D., and Smith, M. (2007). Visualizing the signatures of social roles in online discussion groups. *Journal of Social Structure (JoSS)*, 8(2) :8.
- Wen, J.-R., Lao, N., and Ma, W.-Y. (2004). Probabilistic Model for Contextual Retrieval. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '04, pages 57–63. ACM.
- White, R. W., Bennett, P. N., and Dumais, S. T. (2010). Predicting Short-term Interests Using Activity-based Search Context. In *Proceedings of the Conference on Information and Knowledge Management*, CIKM '10, pages 1009–1018. ACM.
- White, R. W. and Dumais, S. T. (2009). Characterizing and Predicting Search Engine Switching Behavior. In *Proceedings of the Conference on Information and Knowledge Management*, CIKM '09, pages 87–96. ACM.
- White, R. W., Ruthven, I., Jose, J. M., and Rijsbergen, C. J. V. (2005). Evaluating implicit feedback models using searcher simulations. *ACM Transactions on Information Systems (TOIS)*, 23(3) :325–361.
- Widiantoro, D. H., Yin, J., Seif, M., Nasr, E., Yang, L., Zacchi, A., and Yen, J. (1999). Alipes : A swift messenger in cyberspace. In *Proceedings of the AAAI Spring Symposium on Intelligent Agents in Cyberspace*, pages 62–67.
- Wielhorski, K. (1994). Teaching remote users how to use electronic information resources. *The Public-Access Computer Systems Review*, 5(4) :5–20.
- Wilson, T. D. (1981). On user studies and information needs. *Journal of Documentation*, 37(1) :3–15.
- Wood, D. J. and Gray, B. (1991). Toward a Comprehensive Theory of Collaboration. *The Journal of Applied Behavioral Science*, 27(2) :139–162.
- Xie, I. and Joo, S. (2012). Factors affecting the selection of search tactics : Tasks, knowledge, process, and systems. *Information Processing & Management (IP&M)*, 48(2) :254–270.

- Xu, Y., Jones, G. J. F., and Wang, B. (2009). Query Dependent Pseudo-relevance Feedback Based on Wikipedia. In *Proceedings of the Annual International SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '09, pages 59–66. ACM.
- Yao, K.-T., Neches, R., Ko, I.-Y., Eleish, R., and Abhinkar, S. (1999). Synchronous and asynchronous collaborative information space analysis tools. In *Proceedings of the International Workshops on Parallel Processing*, ICPPW '99, pages 74–79. IEEE Computer Society.
- Yee, K.-P., Swearingen, K., Li, K., and Hearst, M. (2003). Faceted Metadata for Image Search and Browsing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 401–408. ACM.
- Yu, C. T. and Salton, G. (1976). Precision weighting - an effective automatic indexing method. *Journal of the ACM (JACM)*, 23(1) :76–88.
- Yue, Z., Han, S., He, D., and Jiang, J. (2014). Influences on Query Reformulation in Collaborative Web Search. *IEEE Computer*, 47(3) :46–53.
- Yue, Z. and He, D. (2010). Exploring collaborative information behavior in context : a case study of e-discovery. In *Proceedings of the International Workshop on Collaborative Information Seeking*, CSCW '10. ACM.
- Zeballos, G. S. (1999). Developing Tools for Efficient Collaborative Web Browsing. In *Proceedings of the Extended Abstracts on Human Factors in Computing Systems*, CHI EA '99, pages 77–78. ACM.
- Zhang, J., Ackerman, M. S., and Adamic, L. (2007). Expertise Networks in Online Communities : Structure and Algorithms. In *Proceedings of the International Conference on World Wide Web*, WWW '07, pages 221–230. ACM.
- Zhang, Z.-K., Zhou, T., and Zhang, Y.-C. (2011). Tag-Aware Recommender Systems : A State-of-the-Art Survey. *Journal of Computer Science and Technology (JCST)*, 26(5) :767–777.
- Zipf, G. K. (1949). Human behavior and the principle of least effort. *Addison-Wesley*.
- Zobel, J., Moffat, A., and Ramamohanarao, K. (1998). Inverted Files Versus Signature Files for Text Indexing. *ACM Transactions on Database Systems (ACM TODS)*, 23(4) :453–490.

